Feature Extraction from Remotely Sensed Imagery for Emergency Management and Environmental Assessment



Doctoral Dissertation
Doctoral Program in Urban and Regional Development (31st Cycle)

Feature Extraction from Remotely Sensed Imagery for Emergency Management and Environmental Assessment

By

Constantin Sandu

******

Supervisor:
Prof. Piero Boccardo, Supervisor

Doctoral Examination Committee:
Prof. Marco Gianinetto, Referee, Politecnico di Milano
Prof. Enrico Corrado Borgogno, Referee, Università of Torino
Dr. Mario Gomarasca, Referee, CNR IREA, Milano
Prof. Mauro Lo Brutto, Referee, Università degli studi di Palermo

Politecnico di Torino
2018

Declaration

I hereby declare that, the contents and organization of this dissertation constitute my own original work and does not compromise in any way the rights of third parties, including those relating to the security of personal data.

Constantin Sandu

14th of December 2018

* This dissertation is presented in partial fulfillment of the requirements for **Ph.D. degree** in the Graduate School of Politecnico di Torino (ScuDo).

I would like to dedicate this thesis to my loving parents

Acknowledgment

I would like to acknowledge Prof. Piero Boccardo, for the support and guidance during my Ph. D study and research.

I would also like to thank my colleagues from ITHACA for their support and patients during this period. Special thanks to Silvana and Fabio for helping me out with the work.

Moreover, I would like to thank Sina and his group from the Department of Electronics and Telecommunications of the Politecnico di Torino for the collaboration.

Most importantly, thank you Susi for the love, care, patience and for cheering me up every time I was in a bad mood.

Abstract

It is known that feature extraction from satellite imagery plays a fundamental role for environmental assessment and during the emergency management phase. Rapid mapping before and after catastrophic events intensively employs automatic and/or semi-automatic procedures to define the situation before the event, and to delineate damages and extract information after it. Earth Observation (EO) data with different spatial and spectral resolution is used for this task since the advent of the first satellites designed for this scope. In recent year the availability and the design of new EO programs created and widespread large amount of free to use satellite data. This thesis has two aims, one is to investigate the use of existing free to use data that acquired from recently launched Multispectral Instrument (MSI) carried by Sentinel-2 to extract surface water, the second aim is to explore the usage of very high resolution (VHR) geometric data to perform building footprint extraction and to assess damages to structures after catastrophic events. Rapid mapping procedures often use computer aided photo interpretation (CAPI) that are very time consuming and are done manually by an operator. The study addressed this issue proposing different type of semi-automatic feature extraction procedures that permit to reduce the use of CAPI. The Surface water extraction was performed on Sentinel-2 data using Google Earth Engine, an emerging data infrastructure as a service. This work was done using a combination of normalized indexes and Otsu's an automatic thresholding technique. The second work was carried out in collaboration with the Joint Open Laboratories of TIM and the Telecommunication Department of the Politecnico di Torino. The aim of the study was to construct a convolutional neural network (CNN) able to extract building footprints

from VHR satellite imagery. The final study was the proposal of a new building damage scale tailored for VHR vertical images.

With the first study it is proposed an operational procedure capable to extract surface water features from Sentinel-2 images in a consistent manner worldwide. The second study proposed a novel approach to define building footprints with promising results. Finally, the third work permitted to propose a standard building footprint damage scale which permitted an amelioration of the accuracy of damages detected from vertical imagery. Furthermore, the scale has been accepted by the international working group on satellite emergency mapping (IWG-SEM). All three studies find an application for emergency management specially during the rapid mapping phase. Additionally, the studies find application also for mapping purposes and for environmental assessment. Future directions will try to address and adapt the proposed procedures to different emerging data processing technologies. Moreover, the quality of the methods will be assessed over more case studies areas and with more accurate ground truth data.

Contents

List of Figures

List of Tables

List of Annexes

Chapter 1

Context

1.1 Introduction

Earth Observation (EO) data has been used to better understand our planet thanks to the first Earth orbiting satellites. Remote sensed data plays an important role in many area of study, e.g. forest cover change (Hansen, et al., 2013), fragile marine ecosystems like coral reefs (Mumby, et al., 2004), landslide susceptibility mapping (Lee, 2004), etc.

Most of the works use different types of techniques to extract information from the EO data. With this work it will be explored the different types of feature extraction methods applied to remote sensed satellite data and the different field of application of the results. The study will stress the importance of EO data and its many applications focusing on the usage of emerging techniques used to extract information useful for operational services (e.g. rapid mapping) and for monitoring purposes.

The presented work will focus on feature and information extraction from EO data characterized by different technical features. In the first part of the thesis it will be established the importance of the work and the necessary background to understand the studied thematic and its application. In a second moment the work will present the research field and the gaps that will be filled. Later the study will focus on the data, methods and tools used to solve the gaps and the described problems. The results and the application of the proposed methods will be described in the final section of the work.

Moreover, there will also be a focus on the application of the methods and the possible improvements to the thesis.

The visual description of the thesis is summarized in **Figure 1.** The work will describe feature extraction from medium and very high-resolution datasets and their application for environmental assessment and emergency management. In particular, it will focus on surface water extraction, building footprint extraction and building damage assessment from vertical imagery.

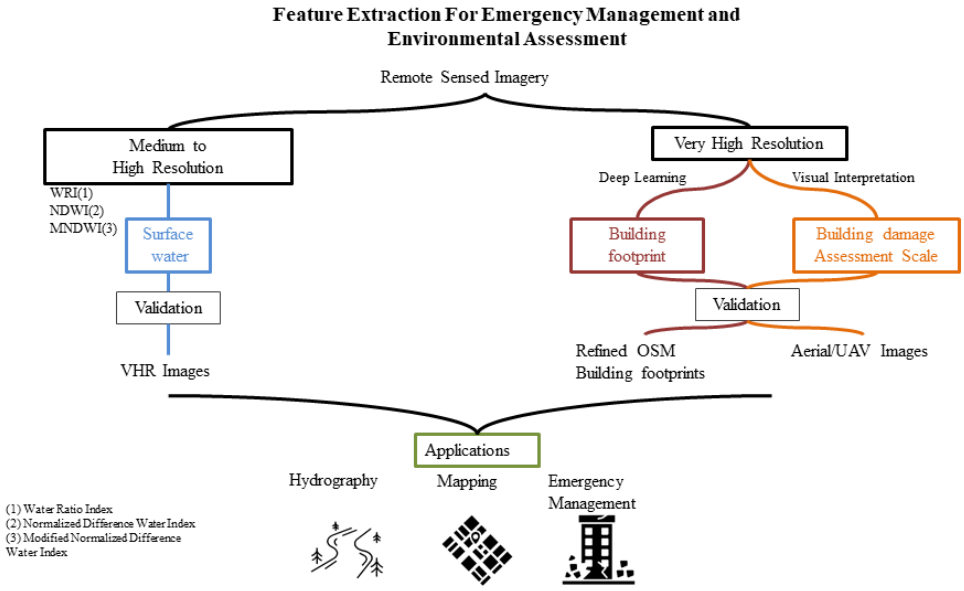


Figure 1 General schema of the thesis

1.2 Natural Disasters

The impact over a community of an extreme natural event turning into a disaster always depends only partly on the force of the natural event itself. The economic and social situation of the region hit by the event is just as significant and determine the country's vulnerability. Humans can only influence to a limited degree weather, and with what intensity, natural events are to occur. However, states can considerable influence the exposure, the susceptibility, the coping and adaptive capacity by their governance in urban planning and disaster mitigation, preparedness and response. This means that states of weak governance are often not able to implement consistent strategies and measures to reduce the disaster risks and consequently their vulnerability is higher. By analyzing the World Risk Index (WRI), which is a mathematical model that combines the physical and spatial exposure to extreme natural events with the societal vulnerability at global scale (Birkmann & Welle, 2016), estimated for 171 countries, it is visible that the

global hotspots of risk due to a high exposure to natural hazards and high vulnerability levels are found in Oceania, Sub-Saharan Africa, Southeast Asia and Central America and the Caribbean — particularly in Haiti.

In terms of vulnerability the states which are facing very high values are mostly locate on the African continent, except for Papua New Guinea, Haiti and Afghanistan. These countries are not only characterized by a high level of persistent poverty but also by significant governance failures and security problems. As example in **Figure 2** is visible that catastrophic events have a huge impact worldwide and cause damages especially in the denser inhabited areas. From 2007 to 2017 catastrophic events have caused 2.040 billion US$ with 611.841 fatalities.

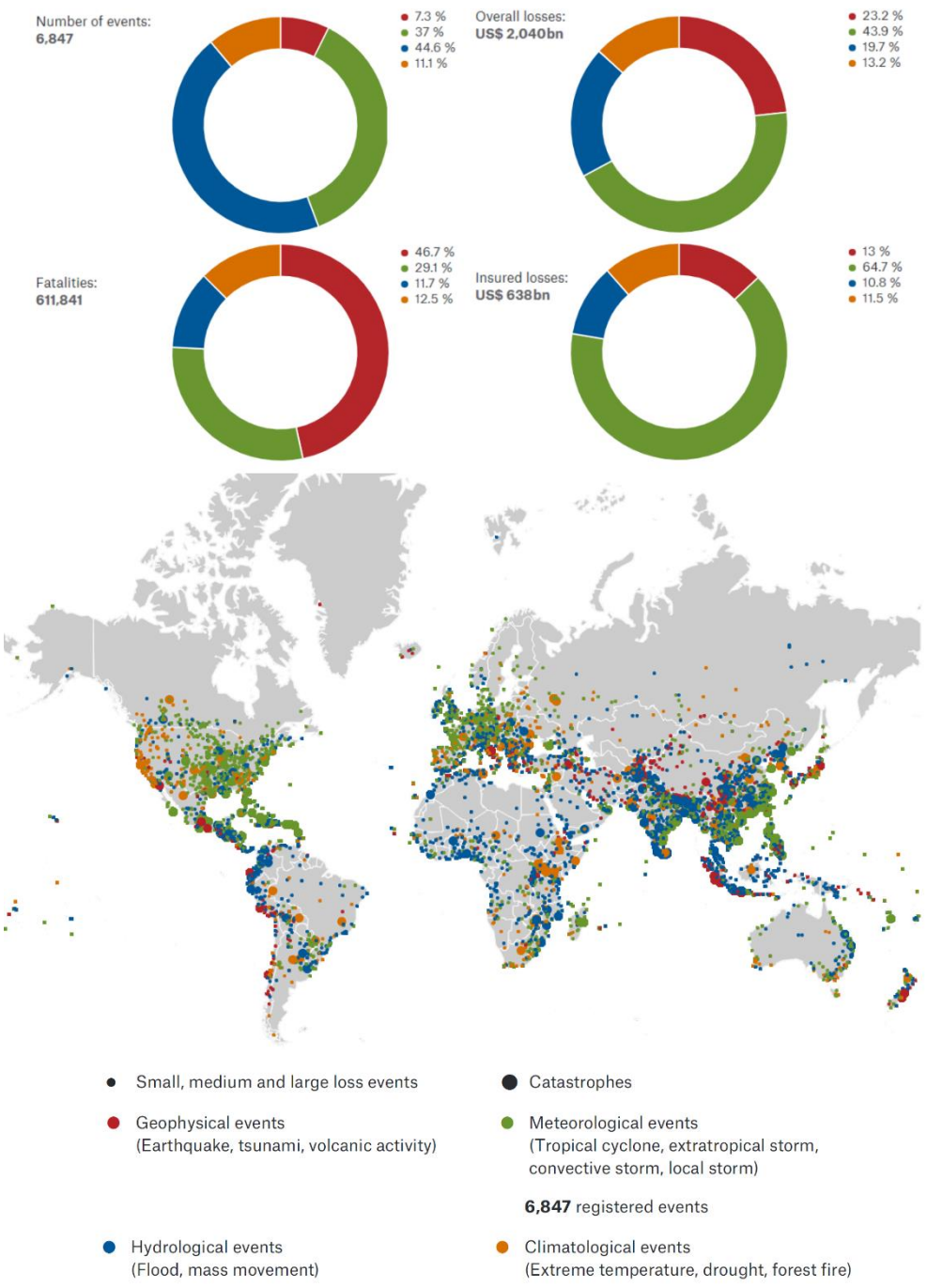


Figure 2 Geographic overview Relevant natural loss events worldwide 2007-2017 (source: Munich Re, NatCatService, 2018)

1.2.1 Economic Impact and Losses

Natural disasters are a fundamental issue in terms of economic losses, affected and killed people. The number of catastrophic events has more than doubled from 1980 to 2011. In 2011 natural disasters caused 27.000 fatalities and approximately 380.000.000.000 US$ in economic losses (Munich Re, 2012).

Major natural disasters can and do have immediate negative short-run economic impacts in terms of structure, infrastructure and tools damages, agriculture and industrial production harms. Disasters also appear to have adverse longer-term consequences for economic growth, social development and food security. While the direct losses are relatively easier to calculate in term of the economic values of the damaged assets, the long-term aftereffects are harder to estimate, and in many cases, especially in under-developing counties, they may be more devastating.

Referring to EM-DAT's (the International Disaster Database maintained by the Centre for Research on the Epidemiology of Disasters (CRED) at the Université Catholique de Louvain in Brussels, Belgium) report of human cost of natural disasters, a total losses of US$ 2,600 billion have been recorded over the period 1994-2013 (Centre for Research on the Epidemiology of Disasters(CRED), 2015). The same report discerns the losses in base on disaster time, geographic location and income groups. In terms of absolute values, Asia accounted for 50% of the total, followed by the Americas at 35% and Europe at 12% (**Figure 3**).

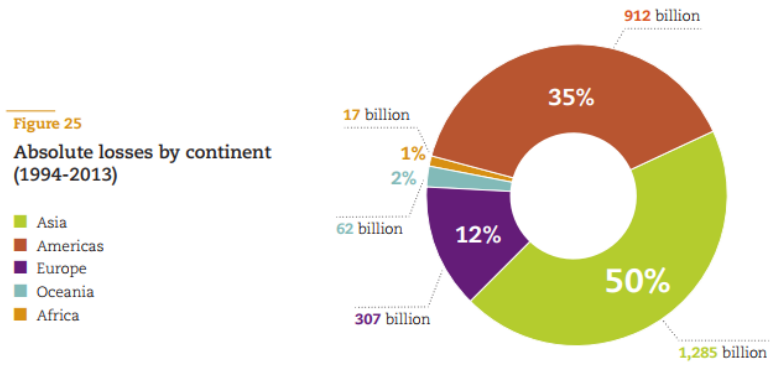


Figure 3 Absolute losses by continent (1994-2013) source: CRED 2015

From the report emerge that storms are the most expensive type of disaster in terms of recorded lost assets (US$ 936 billion), while flood cause more damages on housing worldwide (**Figure 4**).

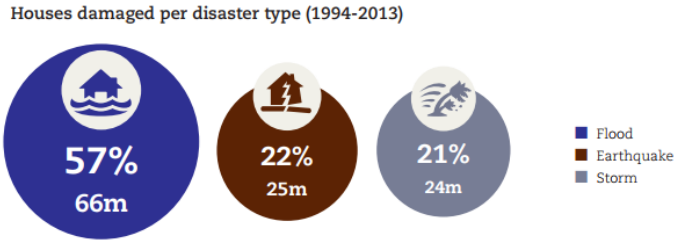


Figure 4 Houses damaged per disaster type (1994-2013) source: CRED 2016

In relation to the income groups low and lower-middle-income countries reported just 10% of economic losses bur more than two-thirds of the deaths, while high- and upper-middle-income countries reported 90% of losses in absolute values and one third of the deaths (**Figure 5**). This discrepancy is not so visible in the case of housing damage, due to the disequilibrium between the higher number of low- and lower-middle-income number of people respect to the lower high- and upper-middle-income number of people, and the value and number of the property.

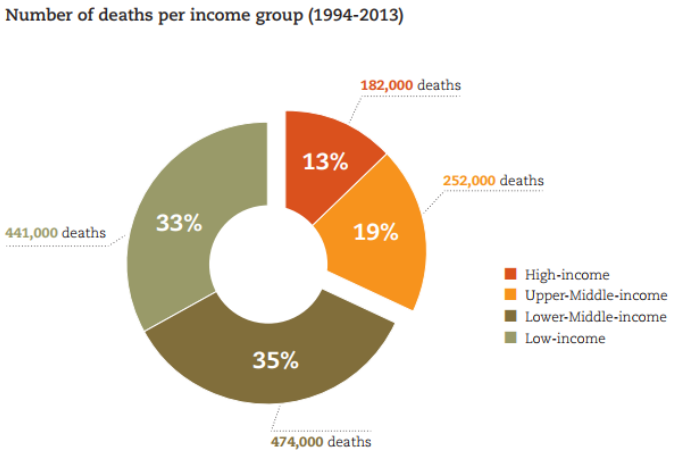


Figure 5 Number of deaths per income group (1994-2013) source: CRED 2016

The impact of natural disasters on national economies also varies greatly at the country level. In absolute values, the USA lost more than any other country between 1994 and 2013, following by Japan and then China (**Figure 6**).

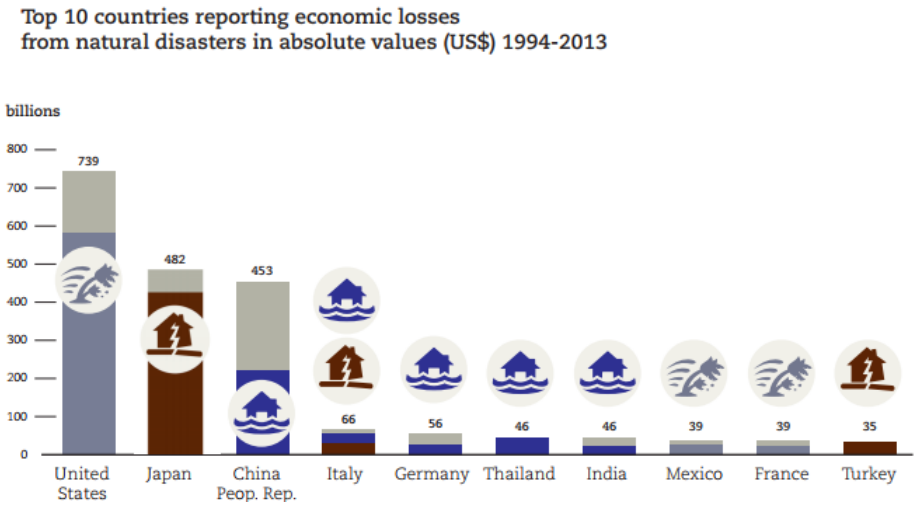


Figure 6 Top 10 countries reporting economic losses from natural disasters in absolute values (US $) 1994-2013 source: CRED 2016

In terms of Gross Domestic Product (GDP), however, the losses were greatest in the Democratic People's Republic of Korea, followed by Mongolia and then Haiti. In fact, despite the higher absolute economic losses in the high and upper-middle income countries, natural disasters cause more harmful impact on the national economy of lower-middle and lower income countries (**Figure 7**).

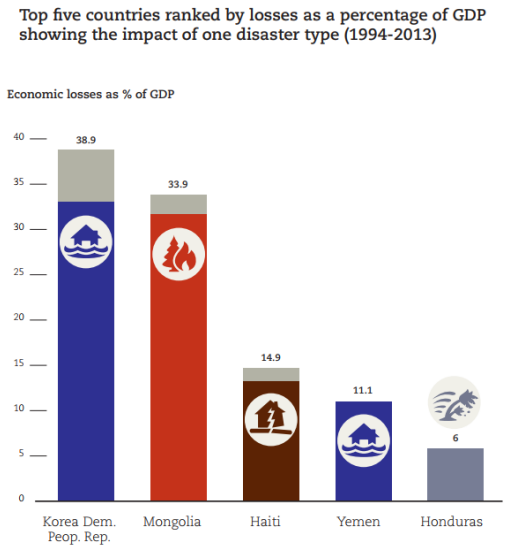


Figure 7 Top five countries ranked by losses as a percentage of GDP showing the impact of one disaster type (1994-2013) source: CRED 2016

Of all-natural hazards, floods, droughts and storms affect the agriculture sector the most, showing the severe impact of climate-related disasters on the sector (Food and Agriculture Organization of the United Nations (FAO), 2015). In terms of the regional distribution of production losses, Asia is the most affected region, with total crop and livestock production losses amounting to USD 28 billion or 40 percent of total losses, followed by Africa with USD 26 billion. In relative terms, Africa is the most affected region, having lost 3.9 percent of total expected crop and livestock production, followed by Central Asia with 3.8 percent (Food and Agriculture Organization of the United Nations (FAO), 2015)

1.2.2 Society Impact

In the period 1994-2013 nearly 2.5 billion people have been affected by natural disaster. Almost 55% of the total are affected by floods, therefore the most deadly events are earthquakes (including tsunamis), which are accounting for 55% of the disaster deaths over the 20-year period, claiming nearly 750,000 lives (Centre for Research on the Epidemiology of Disasters(CRED), 2015) (**Figure 8** and **Figure 9**).

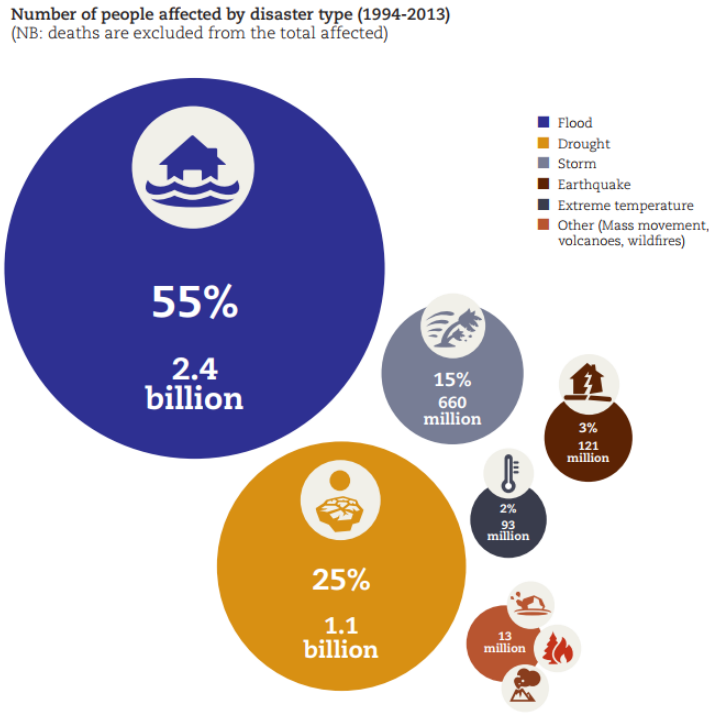


Figure 8 Number of people affected by disaster type (1994-2013) source: CRED 2016

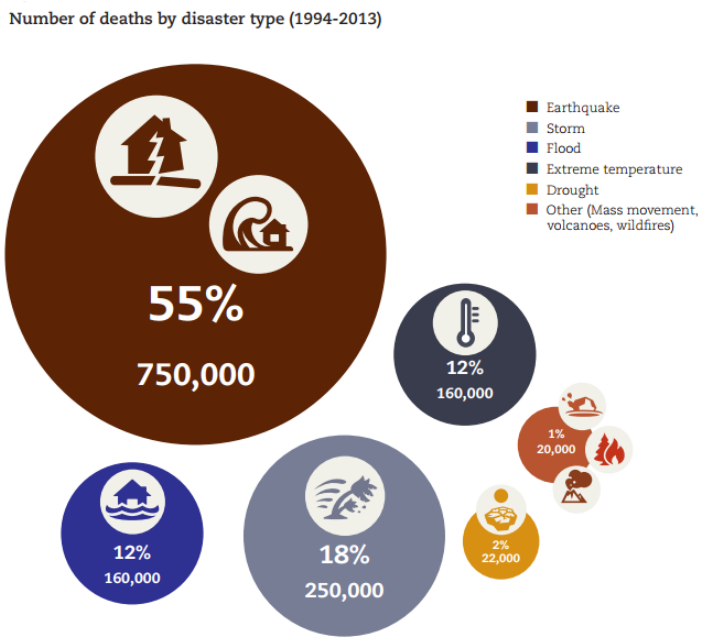


Figure 9 Numbers of deaths by disaster type (1994-2013) source: CRED 2016

The level of people's vulnerability is correlated with social factors such level of poverty, social structure, diversification of income, age and health conditions. Indeed there is a disparity of deaths between lower- and higher-income countries: in the past 20 years high- and upper-middle-income countries experienced 56% of disasters but lost 32% of lives while in the same period, low- and lower-middle-income countries experienced 44% of disasters but suffered a disproportionately high 68% of global mortality (Centre for Research on the Epidemiology of Disasters(CRED), 2015) (**Figure 10** and **Figure 11**).

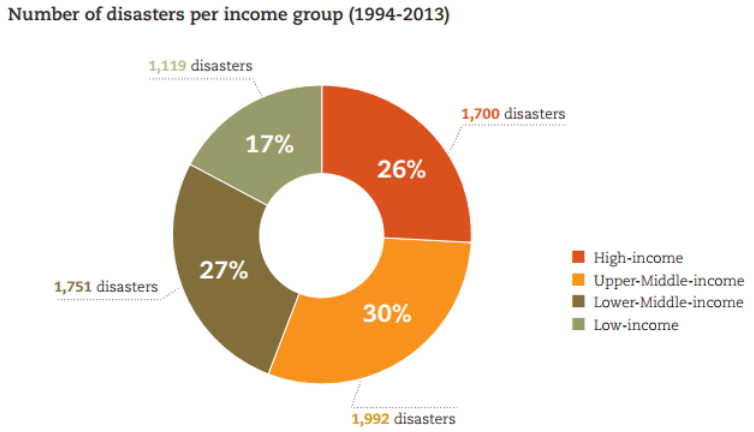


Figure 10 Number of disasters per income group (1994-2013) source: CRED 2016

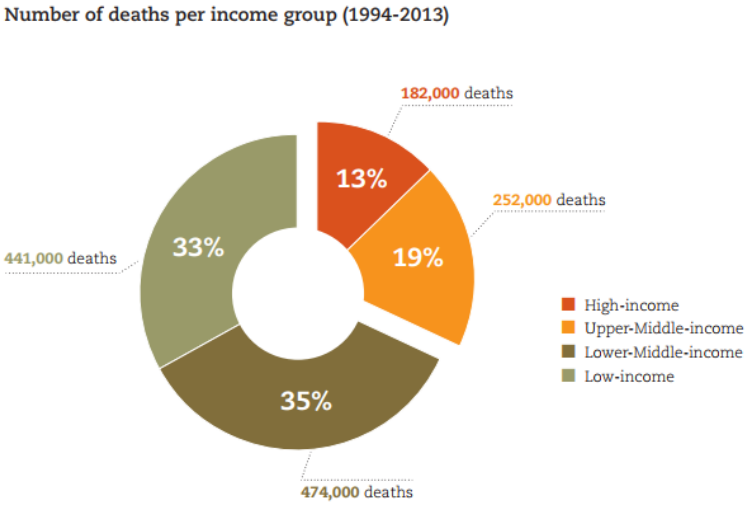


Figure 11 Number of deaths per income group (1994-2013) source: CRED 2016

Natural disasters can have wide-ranging social impacts and can persist for the rest of people's lives. In the long-term, the damages on critical infrastructure (education and health buildings, transportation, power station, water supply etc.), on privet sectors including housing, and agricultural and industrial sectors, can in some cases slow economic and social development for generations, especially in countries where infrastructure were already lacking and people leaved under the threshold of poverty. One of the most immediate effects of natural disasters is the temporal population displacement. When countries are ravaged by earthquakes or other powerful forces of nature, many people have to abandon their homes and seek shelter

in other regions, until normal living conditions have been restored. A large influx of refugees can disrupt accessibility of health care and education, as well as food supplies and clean water. Indeed, disasters have devastating impact on the education systems because of schools being damaged or destroyed, schools being used as evacuation centers, and because disaster risk reduction (DRR) policies are not being adequately resourced or prioritized through different levels of governments and to the community level. Educational inequities may put children at risk of exploitation and violence. Moreover, contributes towards a lack of economic participation (Save the Children, 2016).

In many of the countries more vulnerable to natural disasters, agriculture is the main source of livelihoods and food supply, and a key driver of economic growth. Disasters can challenge the agricultural production, and thus, causing food insecurity. Thousands of people around the world go hungry because of destroyed crops and loss of agricultural supplies. As a result, food prices rise, reducing families' purchasing power and increasing the risk of severe malnutrition or worse.

Trends in migration, as a component of changing population dynamics, have the potential to rise because of natural disaster and extreme climate events, especially when livelihoods are destroyed. Migration can be considered as an adaptation strategy when disasters occur because it helps mitigate the adverse effects on the community by providing new opportunities and resources to the affected people. It is also employed as a coping strategy when other solutions have failed (Mbaye, 2017).

The United Nations Office for the Coordination of Humanitarian Affairs and the Internal Displacement Monitoring Center have estimated that around 20 million people were displaced or evacuated in 2008 because of rapid onset climate-related disasters (OCHA/IDMC/NRC, 2009). Over the last 30 years, twice as many people have been affected by droughts (slow onset events not included in the previous point) as by storms (1.6 billion compared with approximately 718 million) (IOM, 2009). Between 2008 and 2014, developing countries accounted for 95% of global displacement due to disasters. Moreover, most of the displacements are from middle-income countries (**Figure 12**).

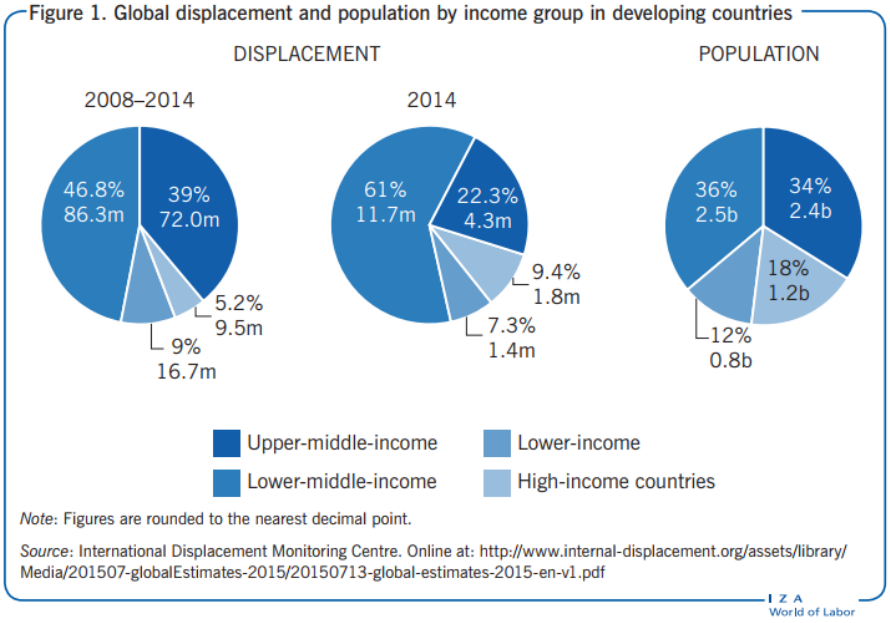


Figure 12 Global displacement and population by income group in developing countries

This is illustrative of the fact that people at the extremities of the income distribution do not necessarily migrate in the aftermath of disasters. Indeed, the poorest cannot afford to migrate and the richest have other mitigation strategies, such as the possibility to recover their lost assets or better access to effective infrastructure and social services, which allow them to cope with disasters without migrating. Therefore, people at the middle of the income distribution are those who do not have many alternatives at their disposal to deal with adverse climatic shocks and, at the same time, can afford migration costs (Mbaye, 2017).

1.2.3 The Role and Benefits of Geospatial Information

In the last 50 years geospatial science has dramatically developed across a variety of sectors and domains. Geographic Information Systems and remote sensing are rapidly becoming more innovative and complex in order to address ever more challenging issues. This technological innovation has produced – and continues to produce - large quantities of data, more than what is used, as already foreseen in 1998 by Al Gore at the California Science Centre:

"The hard part of taking advantage of this flood of geospatial information will be making sense of it. - turning raw data into understandable information. Today, we often find that we have more information than we

know what to do with. (...) Now we have an insatiable hunger for knowledge. Yet a great deal of data remains unused." (Gore, 1998)

An example is the increasing number of satellite sensors, such as the Landsat and Sentinel constellations, which constantly photograph the Earth's Surface with high revisit rates. These records are accessible open to all.

Alongside the explosion of geographical sciences and geomatics, there is an increasing demand for georeferenced information at all levels in society, as it is being accepted as an evidence-based tool, capable of curtailing speculative decision-making. Hence Geospatial information technology is becoming more and more tightly related to political matters. Recently, Geospatial Information Technology has been officially recognized by the United Nations as a necessary tool for monitoring the Sustainable Development Goals at the indicator level and has established a specific Inter-Agency and Expert Working Group on Geospatial Information for the 2030 Agenda on Sustainable Development (United Nations, 2017).

Geospatial Information Technology can be effectively applied to environmental monitoring. The already mentioned significant amount of data with worldwide coverage, the presence a 40-year historical record of satellite-derived data with different spatial and spectral resolutions, smoothly elevates this technology to the preferred tool for monitoring natural phenomena. For instance, oceanographic sciences have leveraged cutting-edge remote sensing technology to analyze bio-physical parameters, such as temperature, chlorophyll concentration and suspended sediments, in oceans and in coastal zones. Ad hoc sensors, characterized by high spectral resolution and wide coverage, such as SeaWiFS, MODIS-Aqua and Sentinel 3 are used by national and international research centers to study Ocean Color and to monitor Harmful Algal Blooms (HABs) (Plymouth marine laboratory, NOAA). A case in point is an EU project launched in 2017, S-3 EUROHAB (https://www.s3eurohab.eu/en/index.html), which addresses the issue of eutrophication and toxic algal blooms in the English Channel, a hazard that is not predictable, and is yet the cause of mass killing of fish and shellfish, summing to a loss of 918 million euros per year to fish industries. As of today, only 3% of the channel is monitored by coastal boat surveys. This project aims at developing a web based Harmful Algal Bloom and Water quality alert system over the entire English Channel, using Copernicus Sentinel-3 satellite, in order to enable marine managers and shell fishermen to monitor the dispersion of such hazardous algae mats.

A special case of environmental monitoring is hazard monitoring and emergency response. In this field remote sensing and geospatial technology has been extensively used, also thanks to the development of Synthetic Aperture Radar (SAR) technologies. For example, observatories constantly keep an eye on geohazard supersites, recording changes in specific variables, through a variety of in situ and remote techniques. The Italian Etnean Observatory, periodically measures ground deformations with SAR interferometry, in order to capture anomalous bulges of the surface, precursor of volcanic eruptions.

Another interesting case is the Open Data Cube initiative, already operational for a few countries (Australia, Switzerland and Colombia). It consists of a method for better managing large sets of satellite data, with the aim of "lowering the technical barriers for users" (https://www.opendatacube.org/about). Specifically, this solution provides analysis ready data with broad temporal and spatial coverage, allowing for studies on land use, vegetation, hydrology and urban planning.

Earth Observations and geospatial information are as of today are irreplaceable tools for the scientific, economic and political world. The abundance of projects and initiatives for improving the use of geoinformation results in the need to standardize the formats and criteria for geospatial data production. International organizations and partnerships, such as the Group on Earth Observations, promote *"open, coordinated and sustained data sharing and infrastructure for better research, policy making, decisions and actions across disciplines"*.

1.3 Remote Sensing and Geospatial Information for Emergency Management

As mentioned previously, geospatial information plays a crucial role for the representation and the study of natural disasters. Satellite emergency mapping (SEM) became more popular along with the diffusion and the availability of a wide range of new remote sensing sensor. As visible in **Figure 13** the number of SEM activations has increased during the past 14 years. In this time initiatives like Copernicus EMS and the United Nations have seen an overall increase use of their services and played a relevant role to produce maps related to catastrophic events.

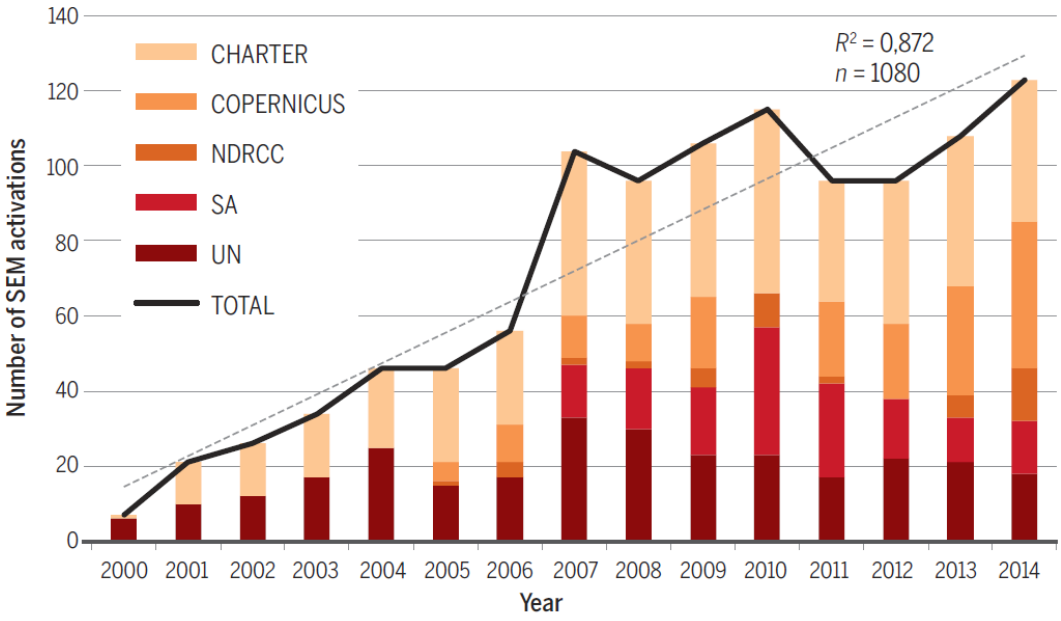


Figure 13 Number of activations and distribution among the different SEM mechanism (source: (Voigt, et al., 2016))

1.3.1 Authoritative Initiatives

International institutions e.g. the United Nations and the European Commission have designed services that provide accurate and up to date geospatial information for the actors involved in the management of natural disasters. The Copernicus Emergency Management Service (CEMS) and UNOSAT in case of major disasters usually provide information about the event very rapidly.

1.3.1.1 Copernicus Emergency Management Service

The Copernicus Emergency Management Service is part of the Copernicus Programme. The Programme is managed by the European Commission (EC) and implemented in partnership with the Member States and the European Space Agency (ESA) the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT), the European Centre for medium-range Weather Forecasts (ECMWF), EU Agencies and Mercator Océan. The Programme is aimed at developing a set of European information services based on satellite Earth Observation and in-situ (non-space) data (European Commission, 2017).

CEMS is directly managed by the European Commission via the Joint Research Centre (JRC) and the General Directorate of the European Civil

Protection and Humanitarian Aid Operation (DG ECHO) provides Early Warning Systems and Emergency Management Mapping Services.

The European Flood Awareness Systems (EFAS) and the European Forest Fire Information System (EFFIS) are the Early Warning Systems managed by the European Commission. EFAS is used to forecast and monitor floods across the Europe. EFFIS provides near real-time and historical information on forest fires in the European, Middle Eastern and North African regions (European Commission, 2017).

The EMS Mapping Service provides geospatial information (raster maps and vector layers in .kmz and ESRI shapefile formats) derived from satellite images. The maps and products that are provided assess the impact of the natural and man-made disaster. The added value of the service is that is 24/7 all year round, supports users all around the world, assesses damages caused by different types of events (droughts, epidemics, humanitarian crisis, infestations, mass movements, storms, volcanic activity, wildfires, forest fires, floods, earthquakes, industrial accidents etc.) (**Figure 14**).

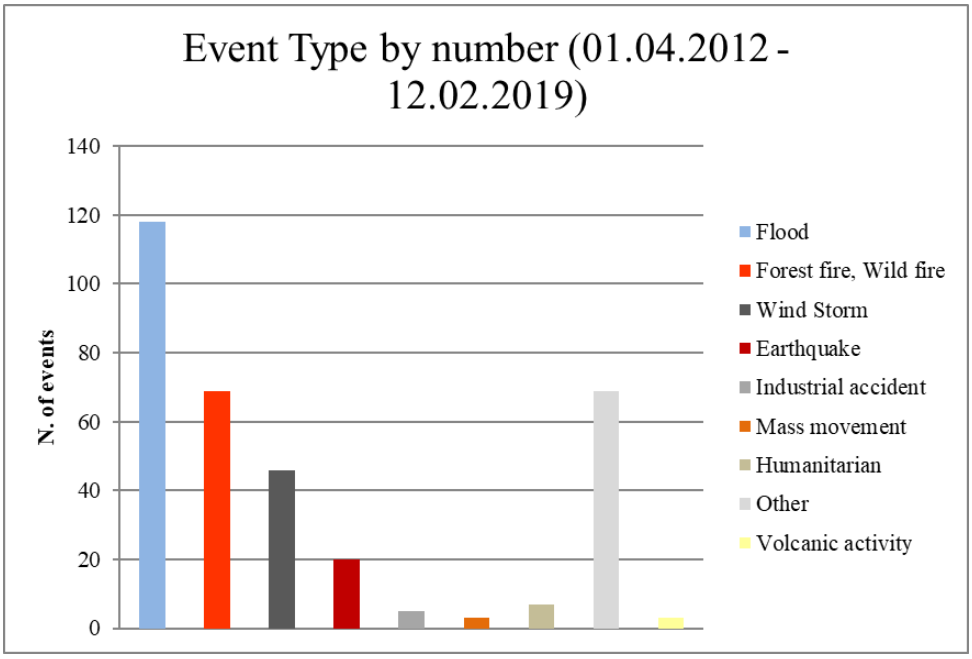


Figure 14 Events analyzed by CEMS

From the moment the service is triggered the service provides reference maps (pre-event information) in 9 hours. The post event maps are provided

after 3 hours after the availability of a post event image and after 12 hours the final map (**Figure 15**).

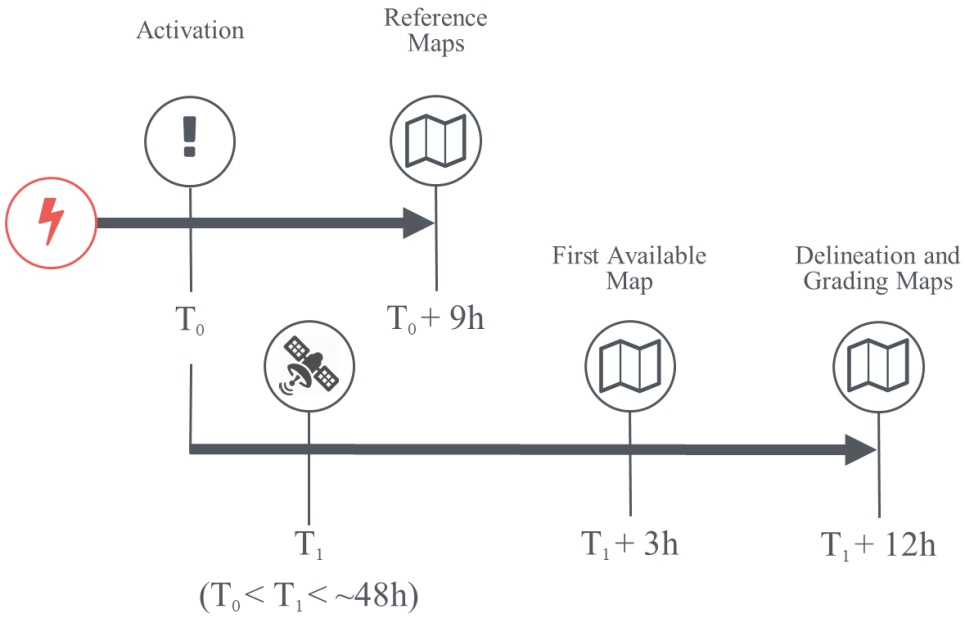


Figure 15 Map production workflow

The three types of products provided by the CEMS is detailed below:

Reference maps: "*provide a quick updated knowledge on the territory and assets using data prior to the disaster. The content consists of selected topographic features on the affected area, in particular exposed assets and other available information that can assist the users in their specific crisis management tasks. A reference map is normally based on a pre-event image captured as close as possible prior to the event"* (Copernicus Emergency Management Service (© European Union, 2012 - 2018)). For example, in case of an earthquake the service provides the building footprints of an affected area prior to the event. Generally, this task is carried out by manual digitalization. Another type of reference layer is the permanent water represented by rivers, canals, lakes and basins that must be up to date. An example can be found here: [EMSR112] Genova: Reference Map.

Delineation maps: *"provide an assessment of the event extent (and of its evolution if requested). Delineation maps are derived from satellite post-disaster images. They vary depending on the disaster type and the delineation*

of the areas impacted by the disaster" (ibid.). The best example is represented by the maps that delineate flood extents and burnt areas perimeters. Relevant examples can be found here: [EMSR314] Lokoja: Delineation Map, [EMSR180] Thassos Island: Delineation Map.

Grading maps: *"provide an assessment of the damage grade (and of its evolution if requested). Grading maps are derived from post-event satellite images. Grading maps include the extent, magnitude or damage grades specific to each disaster type. They may also provide relevant and up-to-date information that is specific to affected population and assets, e.g. settlements, transport networks, industry and utilities" (ibid.).* For example, in case of earthquake the products grade the damages to structures and infrastructures using a tailored damage scale. These maps are used by the ground units to dispatch the rescue units and provide immediate assistance. Examples related to the Central Italy earthquake and other events are visible here: [EMSR177] Capodacqua_Aerial: Grading Map, [EMSR317] Palu: Grading Map, [EMSR320] Port de Paix: Grading Map.

All the described products and all the other maps generated by the CEMS are public and can be viewed and downloaded in raster or vector format from the CEMS portal (https://emergency.copernicus.eu/mapping/list-of-activations-rapid).

1.3.1.2 United Nations Operational Satellite Applications Programme

The United Nations Operational Satellite Applications Programme (UNOSAT) is an organization of the United Nations born in the year 2000. UNOSAT is United Nations Institute for Training and Research (UNITAR) specialized center in Geographic Information Systems and satellite imagery. UNOSAT integrates remote sensing and GIS data through web – mapping and information sharing mechanisms (UNOSAT, 2017).

UNOSAT uses specialized skills to perform satellite analysis, design integrated solutions in GIS and geopositioning, develop the capacity of agencies and recipient countries via training and technical assistance.

The UNOSAT portfolio consists in Maps and Reports that delineate natural or manmade disasters and offer support for the UN or other organizations. The areas of intervention are:

- disaster response;

- humanitarian operations;
- human security and the application of international human law;
- human rights.

Like CEMS, UNOSAT has also a division of rapid mapping in response to major disasters or events that can be monitored using remote sensing and GIS techniques. The Services offered are:

- imagery searching, selection and acquisition;
- image digitizing;
- image analyzing;
- cartographic production;
- report production;
- service publishing and web map production.

The dissemination of the maps is done via standard pdf products and .shp or .gdb for the vector data. Relevant examples of products can be downloaded from the UNOSAT website (http://unitar.org/unosat/maps). UNOSAT has been working lately on providing reports and maps for the floods that affected the Bayelsa State in Niger (http://unosat-maps.web.cern.ch/unosat-maps/NG/FL20180925NGA/Preliminary_Flood_Analysis_Bayelsa_Nigeria_20181011.pdf) and damage the earthquake that struck the Sulawesi area in Indonesia (http://www.unitar.org/unosat/node/44/2856?utm_source=unosat-unitar&utm_medium=rss&utm_campaign=maps).

1.3.2 Voluntary Initiatives

In the past years the World Wide Web had a significant transition. The users no longer only consume content available on the web but also create content. The users are becoming one of the main contributors of the web by sharing knowledge and information (O'Reilly, 2005). Well known examples are Wikipedia the free encyclopedia and other platforms like YouTube used for video sharing and Flickr or Instagram for photo sharing. These platforms use User-Generated Content (UGC) (Carmen & Bucher, 2010) as a main source of information and then display it for the whole internet community.

Volunteered Geographic Information (VGI) can be considered an example of UGC. This type of content is generated by the crowd that uses the WWW. The data can be generated by diverse type of users with different levels of experience in the geodata field (Goodchild M. F., 2007). Significant

initiatives are Wikimapia, Map Creator, Google Map Maker, OpenStreetMap. OpenStreetMap is one of the most famous VGI with over 2.7 million users registered at June 2016.

In a very short time data generated from VGI found usage in diverse services that provide information about catastrophic events. OpenStreetMap data is very easy to download and use. It can be used to produce reference maps about an area. This type map can permit to have the situations before a specific catastrophic event. Moreover, the information can be easily integrated with other geographic information datasets available on the web. Furthermore, the data produced by VGI has seen an increase of its quality in recent years (Zielstra & Zipf, 2010), (Haklay, 2010), (Ali A. L., Schmid, Al-Salman, & Kauppinen, 2014), (Ali & Schmid, 2014) and (Barron, Neis, & Zipf, 2014).

In recent years many services that provide relief in case of natural disasters where born. Two of the most relevant are Humanitarian OpenStreetMap Team (HOT) and Tomnod.

1.3.2.1 Tomnod

Tomnod is a service that was born in 2010 as a research project of the University of California, San Diego and later was acquired by DigitalGlobe (US commercial vendor of space imagery and geospatial content and operator of civilian remote sensing spacecraft). Tomnod's main scope is to analyze satellite imagery to extract information that are useful in case of major disaster or in case of emergencies. This process is a mixture between standard satellite data processing and the support of volunteer mappers. In 2011 Tomnod collaborated with the UNHCR (United Nations High Commissioner for Refugees) to localize the refugees camps present in Somalia, the service was also involved in the mapping of the damages created by the Haiyan typhoon in 2013 and in the search of the MH370 missing plane of the Malaysia Airlines of 2014. Tomnod does not always seek for volunteer mapper only for emergency events, natural calamities or humanitarian crisis, a recent example is the mapping of big portions of Beijing the capital of China, or a recent campaign's goal was to spot and locate the swimming pools of the private properties of Adelaide in Australia (Tomnod, 2016).

The data gathered during these specific crowdmapping activities are not always released to the public. Although, in some cases, generally events that had a huge mediatic impact e.g. the Nepal earthquake of 2015 or the

disappearance of the Malaysian Airlines plane the data can be downloaded from the blog of the company. Another peculiarity of the service is that every action made by the users, even if the users are registered or unregistered, is used to enhance and ameliorate the algorithm that the company uses to pre-classify the satellite images (Tomnod, 2015). Examples that describe the products, which are mainly in vector format can be downloaded from DigitalGlobes Blog. The data generated by Tomnod regarding media covered earthquake like the Ecuador earthquake from 2016 which provides damage grading of buildings can be viewed and downloaded here (http://blog.digitalglobe.com/news/open-imagery-and-data-to-support-ecuador-earthquake-response/) (**Figure 16**). The data is presented in standard formats (shapefiles or geodatabases), but not all data generated by Tomnods campaigns are downloadable and processable.

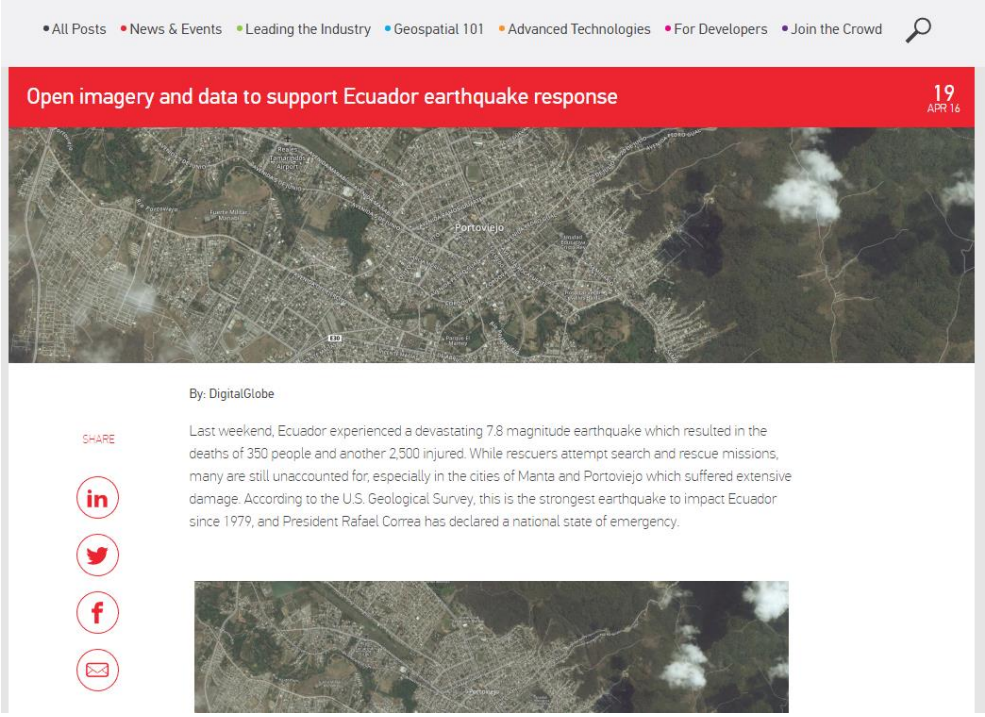


Figure 16 Example of the DigitalGlobe Blog website (source: http://blog.digitalglobe.com/news/open-imagery-and-data-to-support-ecuador-earthquake-response/)

1.3.2.1 Humanitarian OpenStreetmap Team (HOT)

HOT-OSM is a community build on the principles of participation and sharing of geographic data. HOT-OSM and Tomnod share the same purpose.

Although HOT is more focused on the creation of pre-event base data used as support for the creation of reference maps. HOT coordinates the volunteer mappers through an application named Tasking Manager (http://tasks.hotosm.org/). This application permits real time collaborative mapping. The area where the event occurred is divided in smaller zones that can be completed rapidly. Moreover, the Tasking Manager gives in real time the progress of the data generation. To date HOT coordinated and permitted to map more than 1.700.000 tasks relative to diverse type of events and a heterogeneous type of features like roads, buildings, land cover usage, water resources and idp camps (Humanitarian OpenStreetMap, 2013). Al data generated from the HOT community is available under the Open Data Commons Database License(ODbL) (Open Data Commons, 2017). The data can be freely used and shared by the web users and also by external services.

The current state of mapping tasks is visible on HOT OSMs web page (https://tasks.hotosm.org/contribute?difficulty=ALL) (**Figure 17**). An area of interest is delimitated for each task and each area is divided on smaller subtasks. Once a task has been completed, the digitizing work can be validated by a community member (**Figure 18**).

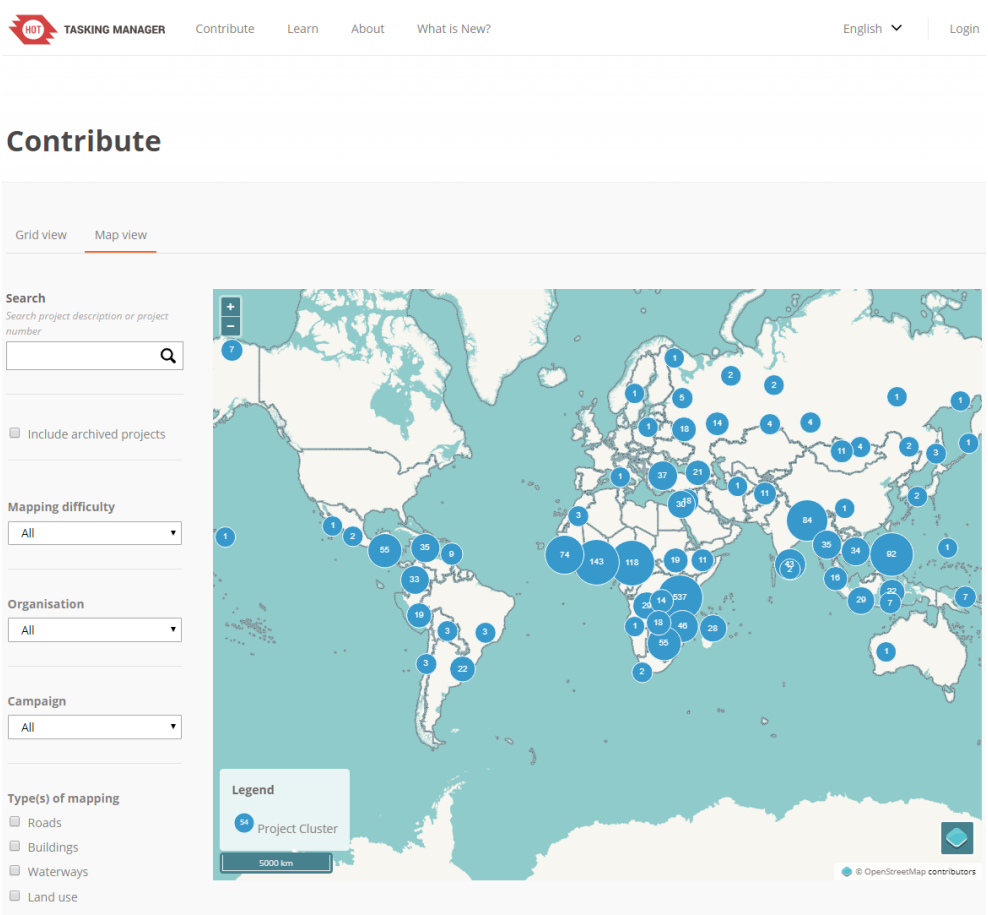


Figure 17 HOT OSM Mapping Tasks (source: https://tasks.hotosm.org/contribute?difficulty=ALL)

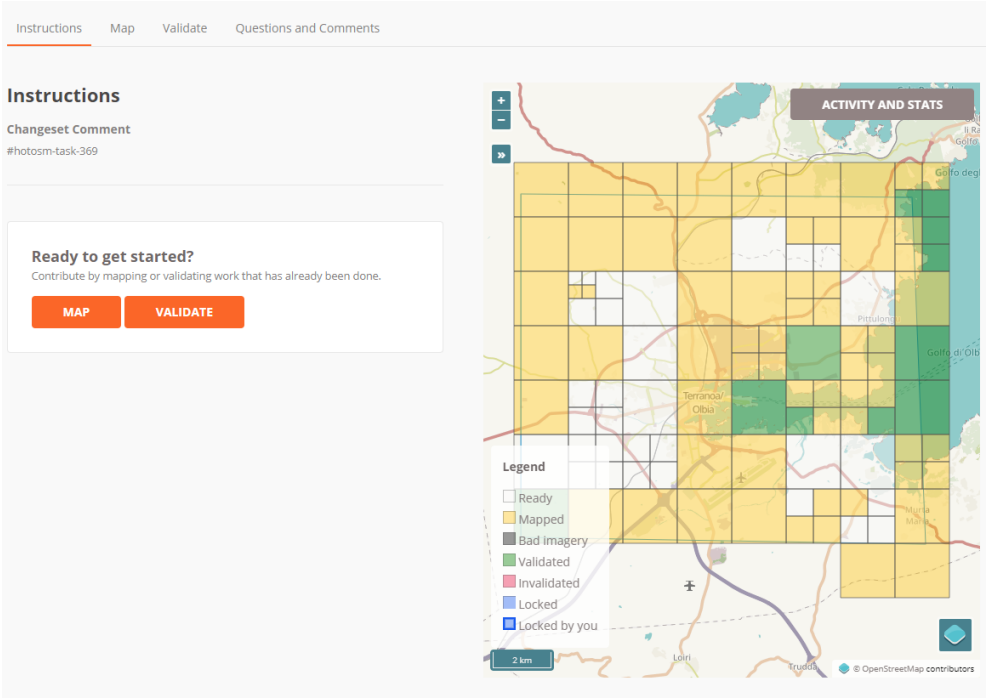


Figure 18 Example of HOT OSM Task over the area of Olbia in Sardinia (source: https://tasks.hotosm.org/project/369#bottom)

Along with these service, it is possible to find data that is produced following rigorous methods and is also free to use. Services like GADM (Global Administrative Areas) and GAUL (Global Administrative Unit Layers) provide the administrative boundaries of the entire globe. Moreover, land cover data of the entire earth surface is available from ESA (GlobCover) and Corine Land Cover.

1.4 Monitoring Services

1.4.1 Copernicus Global Land Monitoring Service

The Copernicus Global Land Service is part of the Copernicus Programme, the EU Earth Observation and monitoring programme. The programme was established by Regulation (EU) No 377/2014 of the European Parliament and of the Council, on April 3rd, 2014. The Copernicus Pogramme (EU Earth Observation and Monitoring programme) has his origins in the 'Baveno Manifesto', (1998, Baveno/Lago Maggiore, Italy).

The Copernicus Global Land Service is the result of several years of implementation. The first monitoring was known as Global Monitoring for Environment and Security (GMES). The programme had an Initial

Operational phase (GIO). The final overall scope, priorities and architecture of the Global Land Service are described and approved during the 33rd International Conference on Remote Sensing of Environment (ISRE33) of Stresa (Italy) in 2009 and revised by the EU Member States Delegates the same year. The global Land Service is designed to answer the needs of the following EU Policy areas (Copernicus Europe's eyes on Earth, 2017):

- Agriculture and Food Security;
- Biodiversity;
- Climate change;
- Forest and Water resources;
- Land degradation & desertification;
- Rural development.

The Copernicus Global Land Service is part of Copernicus programme for the monitoring of the Earth. The Land Monitoring Core Service (LMCS). Thought the processing of Earth Observation data the Copernicus GLS provides added-value products and information about the land surface. This information is provided continually and is delivered to the decision makers and the broad range of users that have direct interest in this thematic.

The service that covers the entire globe and specific areas is a component that is managed by the European Commission’s DG JRC. The data with a higher geometrical and spatial sampling useful at a Local or at a Pan-European scale is a component managed by the European Environment Agency (EEA) (Copernicus Europe's eyes on Earth, 2017).

The Global Land Service managed by the JRC is further organized as follows:

- systematic monitoring at global scale based on low-to-medium spatial resolution sensors and including the constitution of long term, consistent time series;
- hot spot mapping and validation, which is actionable by request and has a limited geographical coverage. The analysis is made on specific regions using high resolution satellite data that has a low revisit frequency;
- ground – based Observations (GBOC), that supports the validation of the products generated by the previous 2 components and also by other services and localized request.

The overall structure of the monitoring programme is summarized in **Figure 19**.

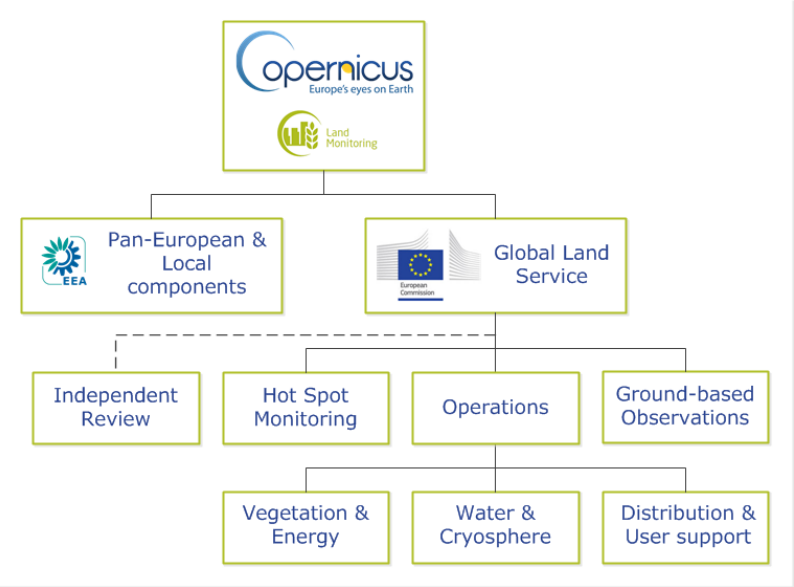


Figure 19 CLMS organization (source: http://land.copernicus.eu/global/about)

Notable examples and projects are Urban Atlas 2012, the project aims to generate a detailed land cover and land use of the major cities of the EU. The data can be viewed and downloaded from the projects web portal (https://land.copernicus.eu/local/urban-atlas/urban-atlas-2012). The data and metadata are described using the INSPIRE format. Riparian zones (transitional zones that occur between freshwater ecosystems and land) (https://land.copernicus.eu/local/riparian-zones/view) and Natura 2000 zones (https://land.copernicus.eu/local/natura/natura-2000-2012)are also mapped in this framework.

1.4.2 North American Land Change Monitoring System

(The North American Land Change Monitoring System (NALCMS) is a joint project between Natural Resources Canada (NRCan), Canada Center for Remote Sensing (CCRS); the United States Geological Survey (USGS); and three Mexican organizations: the National Institute of Statistics and Geography (Instituto Nacional de Estadistica y Geografia – INEGI); National commission for the Knowledge and Use of Biodiversity (Comision Nacional para el Conocimiento y Uso de la Biodiversidad - CONABIO); and the National Forestry commission of Mexico (Comision Nacional Forestal - CONAFOR). The project is facilitated by the Commission for Environmental Cooperation (CEC). The CEC is an international organization created under

the North American Agreement on Environmental Cooperation (NAAEC) by Canada, Mexico and the United States to address regional environmental concerns, help prevent potential trade and environmental conflicts and promote the effective enforcement of environmental law.

The objective of NALCMS (**Figure 20**) is to devise, through a collective effort, a harmonized multi-scale land cover monitoring framework, which can be applied across North America with high accuracy meeting each country specific requirements. The new 2005 Land Cover Database of North America at 250 m spatial resolution is the first step toward achieving this objective.

The product is based on observations acquired by the Moderate Resolution Imaging Spectroradiometer (MODIS). Mapping was performed by each country using unique data pre-processing and information extraction methodologies. These national products were subsequently used to assemble an integrated North America land cover database. The classification legend is designed in three hierarchical levels using the Food and Agriculture Organization (FOA) Land Classification System. Level 1 and 2 are common for North America while level 3 is country specific.

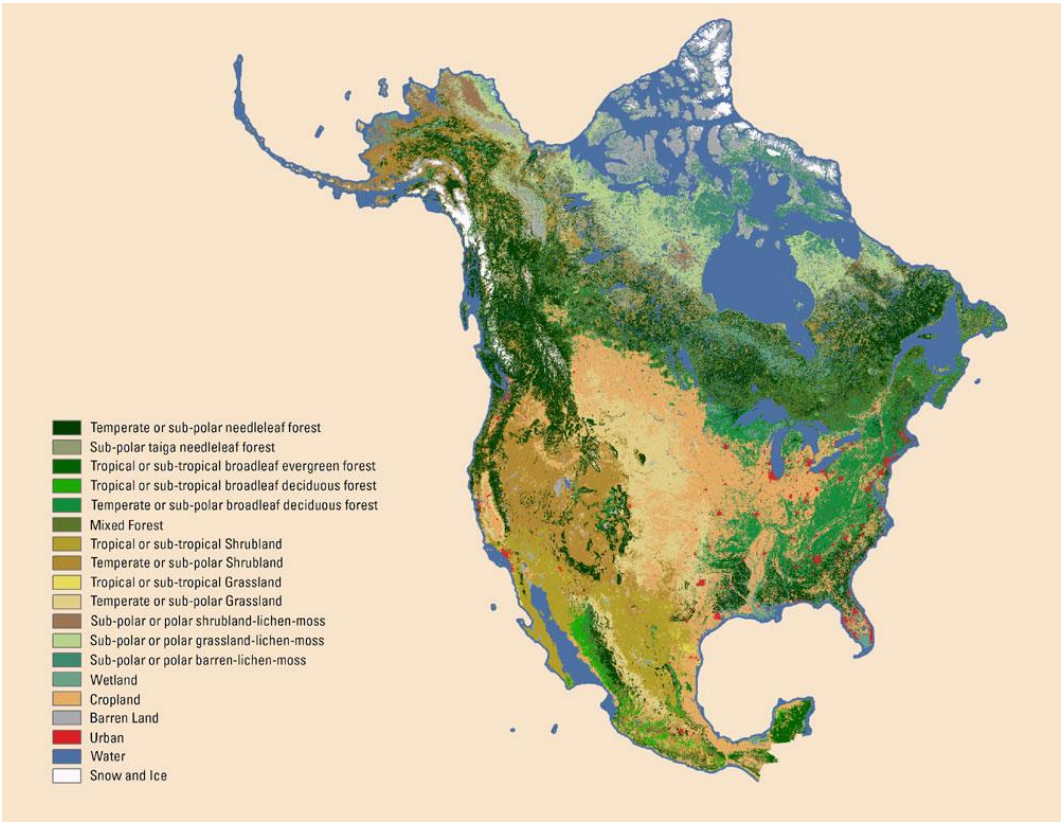


Figure 20 NALCMS Level 1 (source: https://landcover.usgs.gov/nalcms.php)

Information provided by this land cover database is valuable for a range of users including, international organizations such as the United Nations Environmental Programme, non – governmental conservation organizations, land managers, and scientific researchers. The continental scale land cover data generated under NALCMS can be used to address issues such as climate change, carbon sequestration, biodiversity loss, and changes in ecosystem structure and functional, by helping users to better understand dynamics and continental – scale patterns of North America's changing environment.) (United States Geological Survey (USGS), 2016)

Chapter 2

Remote sensed data and its application

2.1 Technical features of the remote sensed images

The main technical features of the remotely sensed datasets will be described in this section. This section will describe features as geometric resolution, spectral resolution and temporal resolution. The combination between the different types of resolution define the characteristics of a satellite. Due to technical constrains, satellites can offer only high spatial resolution with low spectral resolution. On the other hand, satellites with low spatial resolution have higher spectral resolution. Satellite manufactures and space agencies to ensure a higher revisit time build more satellites that orbit opposite to each other (Gomarasca, 2009). E.g. the Sentinel-2 mission has twin polar-orbiting satellites in the same orbit phased at 180° to each other.

Remote sensing has different types of applications. The most common applications are Land use and Land cover, Agriculture, Forestry, Geology, Geomorphology, Damage assessment after catastrophic events, etc. In (GISGeography, 2017) are listed 100 remote sensing applications and uses.

Satellite sensors can be divided in two types, passive sensors and active sensors. *"Passive sensors are microwave instruments designed to receive and to measure natural emissions produce by constituents on the Earth's surface and its atmosphere"* (NASA, 2017). These sensors rely on an external electromagnetic source (e.g. the Sun). "*An active sensor is a radar instrument used for measuring signals transmitted by the sensor that were reflected, refracted or scattered by the Earth's surface or its atmosphere*" (NASA, 2017).

2.1.1 Geometric resolution

These types of analysis can be made because there is a broad availability of datasets generated with different techniques and from different types of platforms. The most common platforms are listed below:

- Satellite Sensor Platforms: Low (MODIS with 500 m ground sample distance (GSD), MERIS 1000 m GSD, Meteosat 3km) and Medium (RapidEye 6 m GSD, Planet 3 m GSD) geometric resolution sensor can be used for Land use and Land cover application. High (Deimos HR 1 m GSD) and Very High Resolution (VHR) (World View 2/3/4, GeoEye can reach VHR 0.5 m and 0.3m GSD) can be used for intelligence applications, security management, and damage detection after catastrophic events;
- Airborne Senor Platforms: Very High resolutions (0.05 m GDS) that can be used to update cartography, delineate catastrophic events that occur at a local scale;
- Unmanned Airborne Vehicles: have similar characteristics as the previous type of vehicles. Moreover, the aircraft doesn't have a pilot on board and can be used to survey areas that can bring harm to the pilot of the vehicle.

The characterization of the sensor can vary based on the spatial, spectral, radiometric and temporal resolution.

The spatial resolution indicates how much 1 pixel of the satellite/platform covers the earth surface (**Figure 21**).

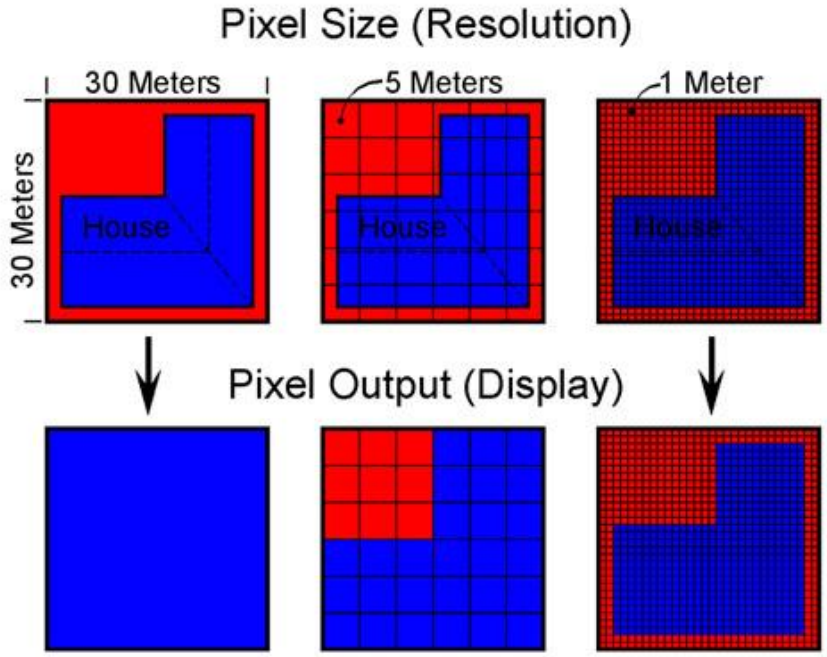


Figure 21 Spatial resolution (source: https://www.satimagingcorp.com/)

- Very high spatial resolution: 0.3 – 1.0 m GSD;
- High spatial resolution: 1.0 – 4.0 m GSD;
- Medium Spatial resolution: 4.0 – 6 m GSD;
- Low spatial resolution: > 30.0 m GSD.

2.1.2 Spectral resolution

The spectral resolution *"describes the ability of a sensor to define fine wavelength interval* (Government of Canada, 2015) *s"*. More intervals equal to a higher resolution (**Figure 22**):

- High spectral resolution: - 220 bands;
- Medium spectral resolution: 3 - 15 bands;
- Low spectral resolution: - 3 bands.

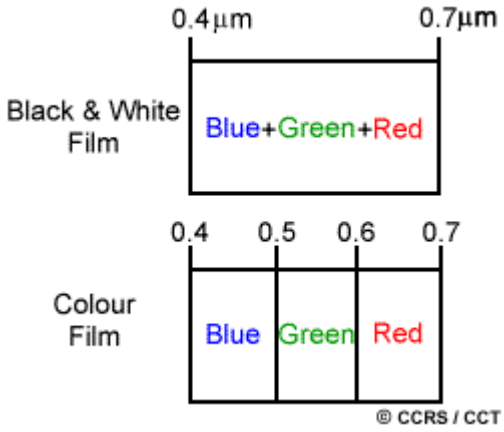


Figure 22 Spectral Resolution (source: http://www.nrcan.gc.ca/node/9393)

2.1.3 Temporal resolution

The temporal resolution of a sensor indicates the revisit frequency of the platform over a specific location:

- High temporal resolution: < 24 hours – 3 days;
- Medium temporal resolution: 4 – 16 days;
- Low temporal resolution: > 16 days;

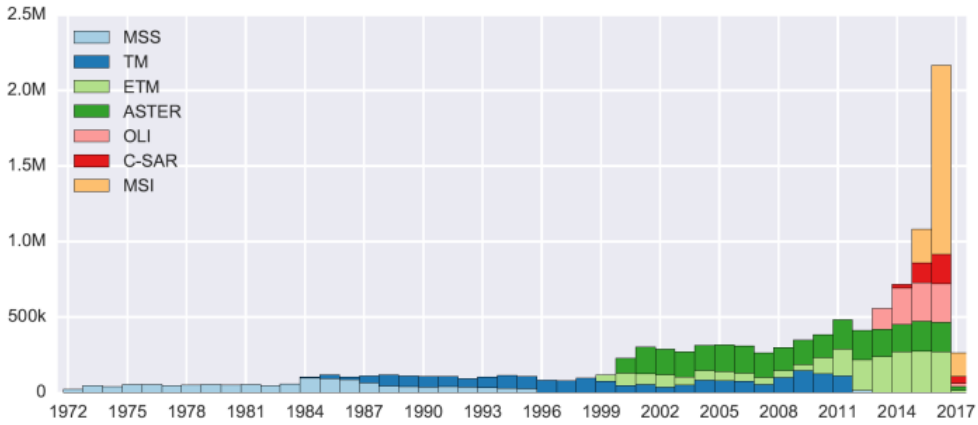


Figure 23 Number or scenes acquired by different satellite sensors from (1972-2017) and available freely source: (Donchyts, van de Giesen, & Gorelick, 2017)

An example of high temporal resolution is visible in **Figure 23**, Sentinel-2 MSI sensor in 2016 has done more acquisition than Landsat ETM, OLI, C-SAR and ASTER sensors combined.

Virtual constellations also play an important role to increase the temporal resolution. Sensors with same resolution permit to have a higher temporal

resolution because virtually the number of scenes over a specific area of interest is higher with two or more sensors in combination rather than one.

Another way to increase the temporal resolution is to re-orient the sensors position during its line of flight. Permitting to the sensor to acquire more scenes from different positions and in different time periods. This is a characteristic of active sensors e.g. COSMO-SkyMed.

2.1.4 Radiometric resolution

Radiometric resolution represents the sensitivity of a sensor to the electromagnetic energy, therefore the reflectance that can be distinguished by the system. *"The radiometric resolution of an imaging system describes its ability to discriminate very slight differences in energy The finer the radiometric resolution of a sensor, the more sensitive it is to detecting small differences in reflected or emitted energy"* (Government of Canada, 2015).

Figure 24 Example of radiometric resolution (source: https://gisgeography.com/bit-depth/)

As visible from **Figure 24** the detail achieved by a sensor with higher resolution at the same GSD have a pixel depth that permits to better define the details of remote sensed images.

2.2 Space missions and applications

2.2.1 USGS Missions

2.2.1.1 MODIS

MODIS is the scientific instrument payload of the satellites Terra (EOS AM) and Aqua (EOS PM) lunched in orbit by NASA in 1999 and 2002. The orbit of the sensor is 705 km above Earth surface with a swath of 2330 km (cross track) by 10 km (along track at nadir). The spatial resolution varies depending on the bands:

- 250 m bands 1-2 (Land/Cloud/Aerosols Boundaries);
- 500 m bands 3-7 (Land/Cloud/Aerosols Properties);
- 1000 m bands 8-36 (Ocean Color/Phytoplankton/Biogeochemistry/ Atmospheric Water Vapor/Surface Cloud Temperature/Atmospheric Temperature/Cirrus Clouds Water Vapor/Cloud Properties/Ozone/ Surface/Cloud Temperature and Cloud Top Altitude).

The data products downloadable from the MODIS web page are divided in 4 groups:

- Level 1 – Raw Radiance, Calibrated Radiance and the Geolocation Fields from the Instantaneous Field of View (IFOV) of the sensor;
- Atmosphere Products – Products related to aerosol, clouds and Atmospheric profiles;
- Land Products – Surface reflectance, Land surface temperature, land cover products, vegetation indexes, thermal anomalies, evapotranspiration, water mask, burnt areas;
- Cryosphere Products – Snow cover, sea ice and ice surface temperature;
- Ocean Products – Sea surface temperature, remote sensing reflectance, particulate organic carbon, particulate inorganic carbon, etc.

2.2.1.2 Landsat

"Landsat represents the world's longest continuously acquired collection of space-based moderate-resolution land remote sensing data. Four decades of imagery provides a unique resource for those who work in agriculture, geology, forestry, regional planning, education, mapping, and global change research. Landsat images are also invaluable for emergency response and disaster relief" (United States Geological Survey (USGS), 2017). The most successful missions like Landsat 4 and 5 provided data about the Earth surface from 1982 to 2009. Landsat 5 exceeded the designed life expectancy and became the longest-operating Earth Observation satellite.

Landsat-8 is the last satellite of the family and as payload carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS). The onboard sensors bands have different resolution depending on the wavelength:

- 30 m – Band 1 to 7 (Aerosol, Blue, Green, Red, NIR, SWIR 1, SWIR 2) and Band 9 (Cirrus);
- 15 m – Band 8 (Panchromatic);
- 100 m resampled to 30 m – Band 10 and 11 (Thermal Infrared 1 and 2).

The data products are downloadable using EarthExplorer and GloVis. The processing Level-1 can be divided in 3 main groups:

- L1T and L1TP: Radiometrically calibrated and orthorectified using ground control points and digital elevation model (DEM) data to correct for relief displacement. These are the highest quality Level-1 products suitable for pixel-level time series analysis;
- L1GT: Radiometrically calibrated and with systematic geometric corrections applied using the spacecrafts ephemeris data and DEM data to correct relief displacement;
- L1G and L1GS: radiometrically calibrated and with only systematic geometric corrections applied using the spacecraft ephemeris data;

2.2.2 Sentinel-2 ESA mission and Copernicus Programme

The SENTINEL-2 constellation is composed by two polar orbiting satellites in the same orbit phased at 180° to each other. The mission is

designed to give a high revisit frequency (5 days at the Equator). Land and climate change monitoring are the main objectives of the programme. The main objectives of the mission are:

- systematic global acquisitions of high-resolution, multispectral images allied to a high revisit frequency;
- continuity of multi-spectral imagery provided by the SPOT series of satellites and the USGS LANDSAT Thematic Mapper instrument;
- observation data for the next generation of operational products, such as land-cover maps, land-change detection maps and geophysical variables (European Space Agency (ESA), 2017).

The Multispectral Instrument (MSI) on board of the satellite measures Earth's reflected radiance in 13 spectral bands from the visible/near infrared (VNIR) to the short infrared spectral range (SWIR). The spatial resolution as for the other sensors depends on the wavelength of the band:

- 10 m – Bands 2,3,4 and 8 (Blue, Green, Red and NIR);
- 20 m – Bands 5,6,7 8A, 11 and 12 (Vegetation Red Edge, Narrow NIR, SWIR);
- 60 m – Bands 1, 9 and 10 (Coastal aerosol, Water vapor, SWIR Cirrus).

The current available level of products for download are generated by the ground segment of the mission or via the SENTINEL-2 toolbox.

- Level-1C Top of atmosphere (TOA) reflectance in cartographic geometry;
- Level-2A Bottom of atmosphere (BOA) reflectance in cartographic geometry (prototype product) can be generated only on the user side using the Sentinel-2 Toolbox.

2.2.3 Landsat and Sentinel-2 comparison

The Sentinel-2 mission was planned and built by the EC and ESA to provide multispectral satellite data with higher resolution in comparison to the native 30 m GSD of the Landsat TM and OLI sensors.

Although the Sentinel-2 data has a higher resolution it is visible (**Figure 25**) in how the data can be integrated with the existing Landsat 7 and 8 data.

Therefore, continue the Earth Observation legacy started in 1982 with Landsat 4-5 Thematic Mapper (TM), the first satellite that consistently acquired multispectral imagery of planet Earth with 30 m GSD.

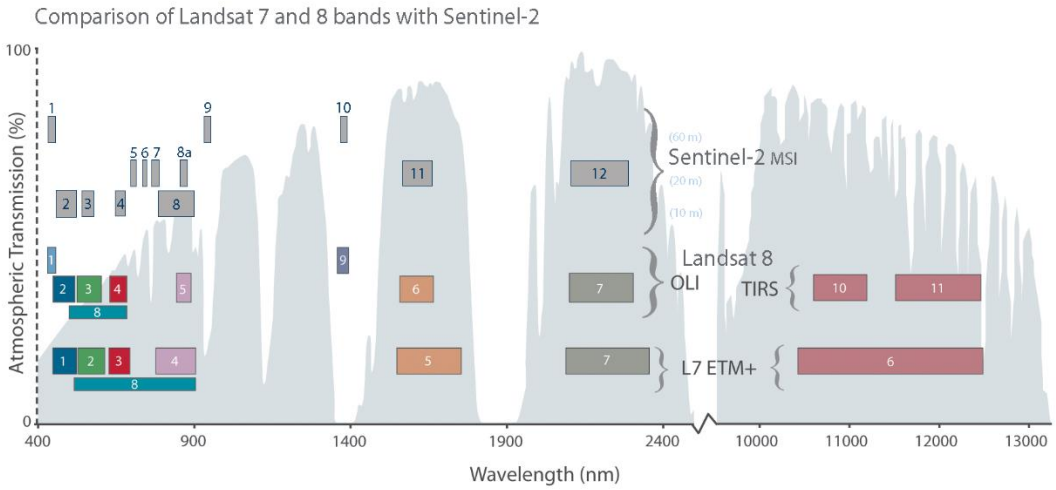


Figure 25 Comparison of Landsat 7 and 8 bands with Sentinel-2 (source: https://landsat.gsfc.nasa.gov/sentinel-2a-launches-our-compliments-our-complements/)

2.2.4 VHR Commercial missions

Since the spring of 1996 and the launch of the first commercially owned and operated satellites (W. Fritz, 1996). The market has seen a rapid growth and demand of VHR satellite imagery for purposes ranging from intelligence application to agricultural mapping. The first commercial satellites provided panchromatic images with a spatial resolution between 1-3 m and a 4-15 m in the multispectral bands.

Today the main commercial data providers (e.g. DigitalGlobe) worldwide are capable to acquire sub metric VHR data from constellations composed by 5 satellites. For example, in the extract of the specification sheet from DigitalGlobe's virtual constellation we see that the panchromatic resolution of the WorldView-3 and the WorldView-4 satellites is 0.31 m GDS (**Figure 26**).

SPECIFICATIONS

Feature	WorldView-1	GeoEye-1	WorldView-2	WorldView-3	WorldView-4
Operational altitude	496 km	681 km	770 km	617 km	617 km
Spectral characteristics	Pan	Pan + 4 MS	Pan + 8 MS	Pan + 8 MS + 8 SWIR	Pan + 4 MS
Panchromatic resolution (nadir)	.50 m	0.41 m	0.46 m	0.31 m	0.31 m
Multispectral resolution (nadir)	N/A	1.64 m	1.85 m	1.24 m	1.24 m
Accuracy Specification (nadir)	6.5 m CE90	3 m CE90	6.5 m CE90	3.5 m CE90	4 m CE90
Swath width	17.7 km	15.3 km	16.4 km	13.2 km	13.1 km

Figure 26 Digital Globe satellite sensor specification (source: https://www.digitalglobe.com/resources/satellite-information)

Airbus DS designed the Pleiades program and projected and build the Pleiades-HR satellites (1A-1B). The satellites have a 0.5 m panchromatic resolution and provide a daily revisit over the same spot of Earth surface.

Therefore, a potential user has the possibility to choose from a virtual constellation of 7 satellites with 0.5 m GSD or lower resolution. This permits to have a wide range of possible providers. The possible virtual constellation increases to more than 15 satellites if users are interested in sub metric data.

2.3 From Data to Information: main processing techniques

2.3.1 Geometric correction, atmospheric correction and radiometric calibration

Remote sensed images usually are not directly usable for environmental studies or applications. Many errors or biases occur during the acquisition of the data, for example misalignment from the planned flight line, suspended particles in the atmosphere, (aerosol) and the conversion from pixel digital

number (DN) to at sensor radiance. Therefore, for most applications it is necessary to correct and calibrate the data before its employment.

Geometric correction is done to integrate acquired satellite data with digital maps or already existing corrected acquisition. Earth's rotation and curvature, together with the flight line of the satellite (that does not always follow the intended orbit) create geometric distortions that must be corrected. The corrections can be done using a DEM and/or ground control points (GCP) with well-known geographic coordinates. An example of geometric corrected data can be viewed in **Figure 27**

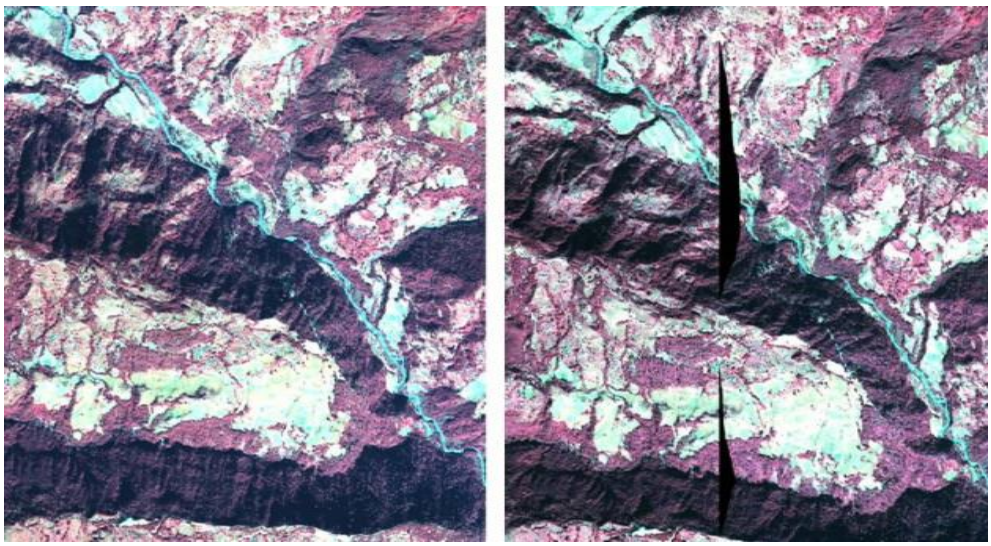


Figure 27 GeoEye-1 image before ortho-rectification (left) and after ortho-rectification (right) (source: https://earthenable.wordpress.com/2015/02/03/guest-post-challenges-geometric-correction-of-optical-high-resolution-satellite-imaging/)

Atmospheric corrections (**Figure 28**) are done to derive an at-ground upwelling radiance (reflectance). This operation is done to correct possible artefacts due to atmospheric scattering. The most common approaches are listed below:

Atmospheric modelling: based on scene specific data (e.g. aerosol content);

Semi-Empirical Modelling: based on the atmospheric visibility and standard constants for date, latitude and longitude;

Empirical Modelling: the data is calibrated using the darkest pixel of the acquired scene (e.g. dark object subtraction).

Figure 28 Example of atmospheric correction with ATCOR (commercial software) (source: https://www.satimagingcorp.com/services/atcor/)

Radiometric calibration is done to convert the DN in radiance and then trough a linear transformation to reflectance. This type of data can be reflectance at sensor (top of atmosphere) or at ground (bottom of atmosphere).

Many data providers release satellite data with different type of processing. As mentioned before, the Sentinel-2 data is released with two different level of service (Level 1C and Level 2A). And for the proposed methods was sufficient the Level 1C of processing. However, the data used for the building footprint extraction and the building damage scale in this work needed to be corrected geometrically, were radiometric calibrated at sensor, and were not atmospherically corrected. This was done to simulate a Rapid Mapping workflow where atmospheric and radiometric calibration are not implemented due to time constraints.

2.3.2 Surface Water Extraction

Surface water extraction is commonly implemented using both passive and active satellite sensors. The datasets generated by MODIS and the Landsat legacy satellites (passive sensors) are the most common and widely used datasets. The datasets are free to use and easily downloadable from the USGS EarthExplorer web platform. Since the end of 2015 ESA developed the Sentinel-2 mission that will provide imagery for the monitoring of the Earth surface. The data is free to exploit and has a higher geometrical accuracy.

To date active sensors have been rarely used in surface water extraction only for monitoring purpose. This is mainly because Synthetic Aperture Radar (SAR) satellites are commercial and the acquired imagery is not free. In 2010 the European Union founded the Sentinel-1 space mission consisting of two satellites, giving continuity at the previous ERS1, ERS2 and Envisat ESA satellites. The mission is carried out by ESA within the Copernicus Programme.

There are other SAR and optical satellites that acquire data about the Earth surface but considering the applications it has been chosen to work using the medium to high resolution images offered by the Landsat and Sentinel missions.

2.3.2.1 Information Extraction Techniques

An extensive review of water bodies classification techniques presented in literature was conducted, to identify those capable to solve or minimize the above-mentioned problem. Since the institution of an historical archive requires the processing of several satellite data, it was decided to investigate on classification procedures that require short processing time. Thus, the literary review was focused on simple classification techniques, mainly based on indexes derived from differential band ratios or histogram thresholding.

The advantage of the indices is that they make threshold values more independent from image acquisition parameters. According to literature two indices are commonly used to identify and classify flooded areas: The Normalized Difference Vegetation Index (NDVI) and the Normalized Differential Water Index (NDWI). The extraction of the flooded areas is mainly done by band ratios and histogram thresholding.

The NDVI is an index generally used for vegetation analysis (see chapter 5), but it can find application also for the detection of water bodies. Its definition is:

$$\mathrm{NDVI} = \frac{\mathrm{RED} - \mathrm{NIR}}{\mathrm{RED} + \mathrm{NIR}}$$

Where RED and NIR stand for the reflectance measurements acquired in the red and near-infrared regions, respectively.

Since water absorbs energy in the IR band, NDVI present low values in correspondence of flooded areas but also in correspondence of bare soils, which have similar characteristics both in the visible than in the IR band. For

this reason, to extract water bodies, NDVI values are often combined with IR histogram thresholding, which allows to mask bare soil areas.

The NDWI is instead generally used for describing temporal and spatial dynamics of surface moisture, but application have been found which use this index combined to IR thresholding. Unfortunately, a unique definition of the NDWI was not found, probably due to its adaptation to the different characteristics of spectral sensors mounted on satellite platform normally used for those applications.

The most diffused definition of NDWI was formulated by (McFeeters, 1996), and later also described by, (Chatterjee, Kumar, Chakravorty, Lohani, & Kumar, 2005), (Jain, Saraf, Goswami, & Ahmad, 2006), is computed using the reflectivity in the green and near-infrared bands:

$$\mathrm{NDWI} = \frac{\mathrm{GREEN} - \mathrm{NIR}}{\mathrm{GREEN} + \mathrm{NIR}}$$

GREEN and NIR stand for the spectral reflectance measurements acquired in the green and near-infrared regions, respectively.

This index reduces commission errors during classification, due to vegetation and bare soil classes. (Xu, 2006), (Fengming, Bing, Huabing, Qian, & Peng, 2008)highlight the low reliability of this index in urban areas, proposing a Modified Normalized Difference Water Index (MNDWI) to minimize also errors due to the presence of shadows:

$$\mathrm{MNDWI} = \frac{\mathrm{GREEN} - \mathrm{SWIR}}{\mathrm{GREEN} + \mathrm{SWIR}}$$

where GREEN and SWIR stand for the spectral reflectance measurements acquired in the green and Short-wavelength infrared regions, respectively.

(Shen & Li, 2010), propose the below WRI definition, for the identification of water bodies:

$$\mathrm{WRI} = \frac{GREEN + RED}{NIR + SWIR}$$

Where NIR and SWIR stand for the spectral reflectance measurements acquired in the near-infrared and Short-wavelength infrared regions, respectively.

Many other indexes that permit to extract water have been proponed below some of the most used and notable ones reviewed in (Huang, Chen, Zhang, & Wu, 2018):

Indices	**Equation**
NDWI	NDWI = (GREEN − NIR) / (GREEN + NIR)
mNDWI	mNDWI = (GREEN − SWIR) / (GREEN + SWIR)
AWEI	AWEInsh: 4 × (GREEN − SWIR1) − (0.25 × NIR + 2.75 × SWIR2) AWEIsh: BLUE + 2.5 × GREEN − 1.5 × (NIR + SWIR1) − 0.25 × SWIR2
WI2015	1.7204 + 171 × GREEN + 3 × RED − 70 × NIR − 45 × SWIR1 − 71 × SWIR2

Table 1 Several Popular water indices along with their equation (Huang, Chen, Zhang, & Wu, 2018)

But for this study we proposed the modified water index (MWI) which is result of the addition of the binary classifications obtained by thresholding WRI, NDWI and MNDWI indexes formulated as follows:

$$MWI = \left(\frac{GREEN + RED}{NIR + SWIR} \Big| > th\right) + \left(\frac{GREEN - NIR}{GREEN + NIR} \Big| > th\right) + \left(\frac{GREEN - SWIR}{GREEN + SWIR} \Big| > th\right)$$

The MWI will be used to compute the extraction of the surface water. The MNDWI performs better than other indexes when we are in presence of shadows. Although, it's not easy to extract surface water that is characterize by the presence of suspended material (e.g. mud or organic matter). By thresholding each synthetic index and later summing them and obtaining a new post classified image, the water bodies are enhanced and the unwanted features that generally generate commission errors (e.g. cloud shadows, topography shadows, building shadows) are omitted.

2.3.2.2 Thresholding technique

In this context the goal is to classify the image pixels in two classes, water and not water. Classifying an image can be done using many different thresholding techniques. One way to find the optimal threshold is by adopting a trial and error method inspecting the pixel values present in the satellite image. This method is straight forward and very effective during time critical applications. Although, trial and error methods are not optimal.

The Otsu method tries to overcome this problem by finding the value that best maximizes the inter-class variance. The formulation of the method is the following:

$$BSS = \sum_{k=1}^{p} (\overline{DN}_k - \overline{DN})^2$$

Interclass variance is formulated as BSS which represents the between sum of squares. The two classes (water/ not water) are represented by *p. DN* is the digital numbers of each normalized difference synthetic band computed. $\overline{DN}_k$ is the mean digital number in *k* class and $\overline{DN}$ is the mean digital number present in the entire dataset. Class k is defined by every DN less than some threshold. The goal is to find the threshold that maximizes the BSS.

This approach will not be done on the whole image. To better optimize the procedure the Otsu method will run only on the bins of a histogram. This approach is fast and requires only one pass over the data. At each bin of the histogram, define class k as the pixels in that bin and lower. Class *k+1* is everything else. The method permits to find the mean that maximizes the BSS. This method is very fast but has to take into account that the sample used to generate the histogram has to contain water and not water DN in roughly equal proportions.

2.3.2.3 Novel water extraction techniques

In recent years different approaches that combine the extraction techniques mentioned before with image segmentation processes where applied. These techniques are mainly applied for the water detection of lakes. For example, in (Donchyts, van de Giesen, & Gorelick, 2017) there is a combination between the NDWI calculation, the segmentation (using Canny edge detection) of the image from whom it has been extracted and finally the

thresholding using the Otsu method over the pixels that fall within a buffer of the edges that have certain characteristics. A similar method is also applied by EO Learn, an open source python library that permits to extract surface water from Sentinl-2 images. The later after the definition of an AOI and a pre-loaded water mask of a waterbody permits to create a task that generates the surface water level of the input waterbody. The method is summarized in **Figure 29** and is divided in the following steps:

1) Area of interest Definition;
2) Load of the satellite imagery (one image or multiple images);
3) Cloud detection;
4) NDWI Calculation;
5) NDWI thresholding (using Otsu's method);
6) Vectorization of the results;
7) Export of the results to a DBMS;

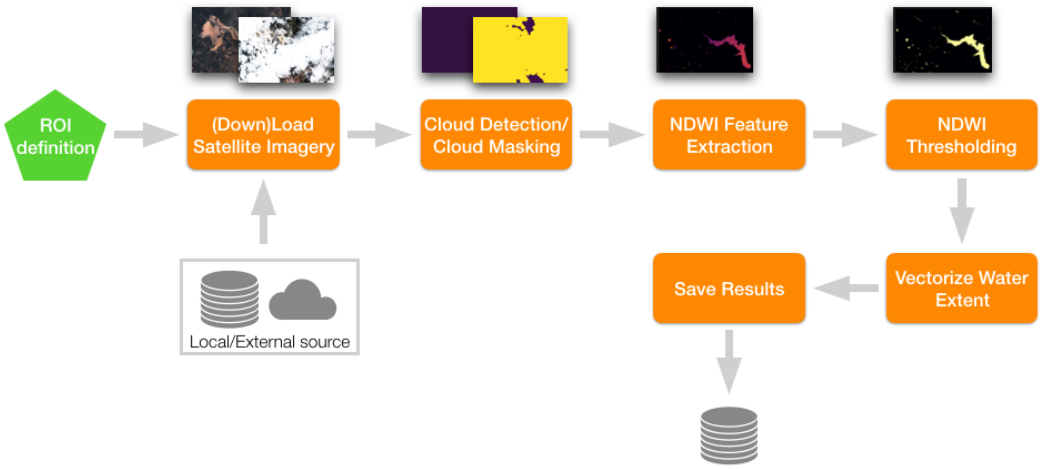


Figure 29 EO-Learn Surface water extraction method (source: https://medium.com/sentinel-hub/introducing-eo-learn-ab37f2869f5c)

2.3.3 Building Footprint Feature Extraction

This section will describe the application of remote sensing for built-up area delimitation using medium to high resolution data. Moreover, the section will also focus on the state of the art of building footprint extraction using VHR satellite images.

2.3.3.1 Information Extraction Techniques

Feature extraction from satellite images has a crucial role not only for emergency management or the definition of the built environment of urban

and rural centers. This activity could also play an important role in the update of the maps and data available on-line.

In recent times there has been a huge interest in this topic and different techniques have been proposed at the different scales and resolutions of the datasets used for the analysis.

2.3.3.2 Low Resolution Datasets

The Normalized Difference Built-Up Index (NDBI) (Zha, Gao, & Ni, 2003) is one of the most used index and technique for mapping urban areas using Landsat Thematic Mapper data (formula below):

$$\mathrm{NDBI} = \frac{TM\ Band\ 5 - TM\ Band\ 4}{TM\ Band\ 5 + TM\ Band\ 4}$$

Moreover, this index is used also with the newer Landsat-8 satellite. The technique combined with other indexes like the Normalized Difference Vegetation Index (NDVI) and the Normalized Difference Water Indexes (NDWI, MDWI) permits to achieve good result for the urban areas feature extraction (Zha, Gao, & Ni, 2003) (Bhatti & Tripathi, 2014) (Sinha, Verma, & Ayele, 2016).

Other approaches propose the usage of indexes that are peculiar for each analyzed scene (Angiuli & Trianni, 2014) and take advantage of new platforms such as Google Earth Engine (Tiranni & al, 2015) (**Figure 30**).

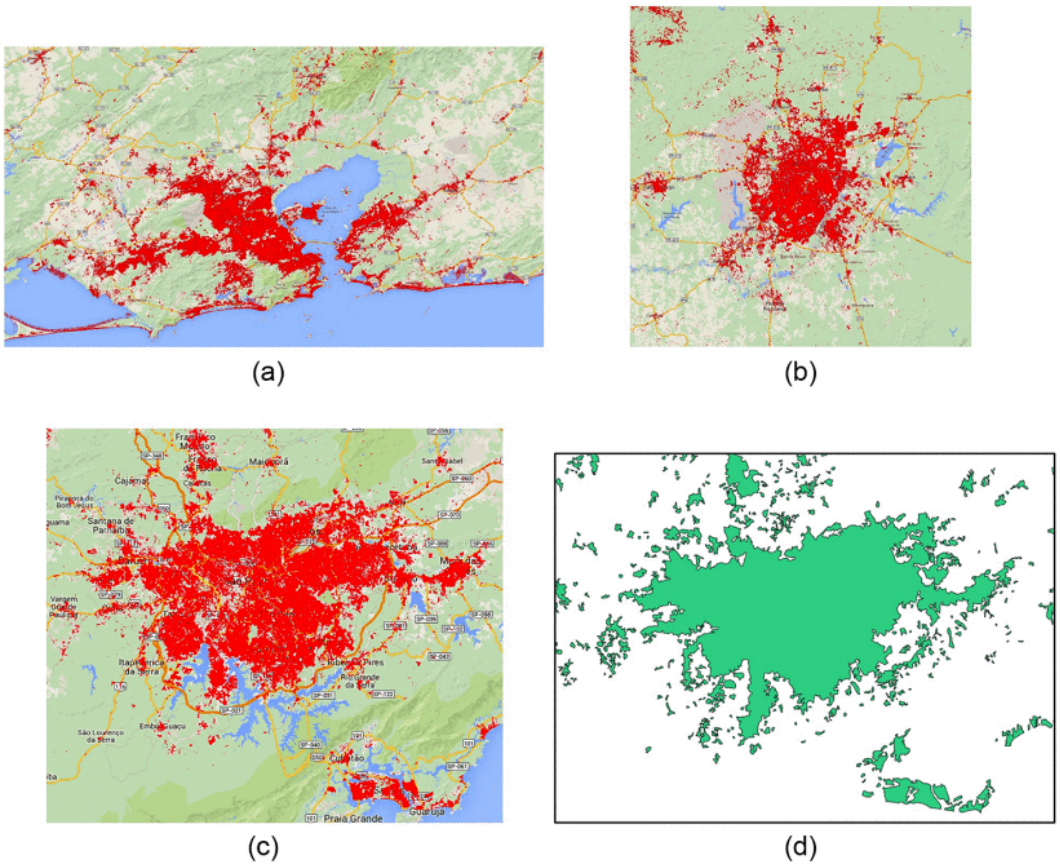


Figure 30 Human settlement extent extraction results (Tiranni & al, 2015)

2.3.3.3 VHR Datasets

This family of procedures is no longer based on the single pixel extraction and classification. Working with very high-resolution datasets is often challenging because there are factors that can influence the correct extraction of features. The off-nadir angle is often one of the most important factors. The IKONOS imagery takes advantage of the off nadir angle and can determine the position and the height of buildings in dense residential areas (Shaker, Abd-Elrahman, K. Abdel-Gawad, & A Sherief, 2011).

Object-Based detection algorithms became more popular and started to be applied on VHR imagery. These type of algorithms take in consideration not only the radiometry of each pixel but also segment the image and look for shapes that can be aggregated in a second moment (Attarzadeh & Momeni, 2012) (Li, Wang, & Li, 2014).

Artificial neural networks are one of the last trend used in remote sensing for building extraction. This technique is inspired from the computing

systems and informatics. These techniques have been applied in different works on low and high resolution imagery (Awad, Chehdi, & Nasri, 2007) (Neagoe & Strugaru, 2008) (Zahra & Hamid, 2017).

2.3.4 Building damage assessment from vertical images

This section will describe the current state of building damage assessment using remote sensed vertical images. Damage assessment activities are mainly done after catastrophic events and are done during rapid mapping activations.

2.3.4.1 Information Extraction Techniques

The main operational steps (and related timeline) of a simplified Rapid Mapping general workflow are shown in Figure 31. To date, standard procedures (**Figure 31**, red box) for structures and infrastructures damage severity assessment based on post-event VHR vertical imagery following natural disasters is generally carried out adopting a multi-temporal approach comparing baseline data (e.g., imagery captured before the event) to post-event imagery (Xinjian & Yin, 2004).

As far as earthquakes are concerned, although several semi-automated approaches (including the exploitation of promising deep-learning algorithms) are currently being tested, photo interpretation (Plaza & al., 2009) is still the most common methodology to rapidly generate earthquake damage assessment as specified in “Rapid Mapping: geomatics role and research opportunities” (Ajmar, Boccardo, Disabato, & Giulio Tonolo, 2015). The same paper also highlights the need to systematically improve the computer aided photo interpretation (CAPI) both in terms of efficiency and thematic accuracy, to increase the reliability of infrastructure damage assessment information and to improve the timeliness of the crisis information delivery, crucial in rapid mapping. The adopted damage scale and the related interpretation guidelines are clearly impacting on both the before mentioned goals. Another key factor affecting the building damage assessment accuracy and level of detail is the type of remote sensing imagery exploited for the analysis, mainly depending on the platform on which the imaging sensor is installed.

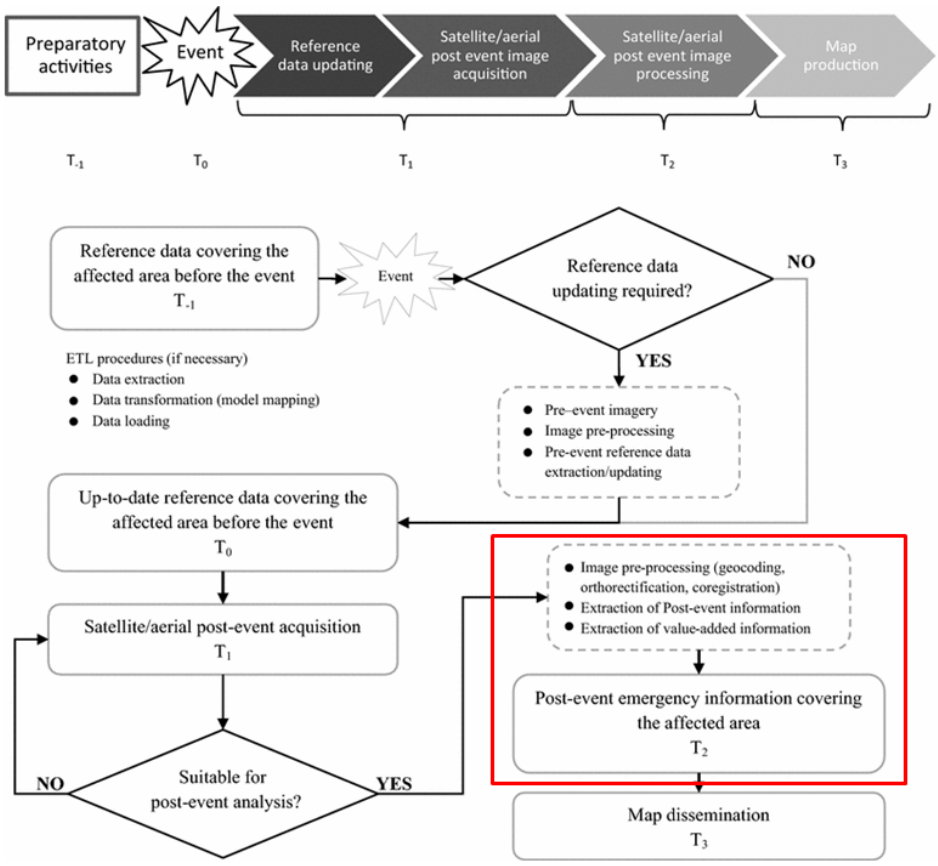


Figure 31 Simplified Rapid Mapping general flow-chart highlighting the main processing steps (crisis information extraction in the red box) and the activity timeline (Ajmar, Boccardo, Disabato, & Giulio Tonolo, 2015)

2.3.5 Accuracy assessment

A ground truth is required to assess the accuracy of a classification datasets and is one of the input required to calculate the confusion matrix, a table that shows correspondence between the results of a classification process and reference data. It is usually used as the quantitative method of characterizing the thematic accuracy of a dataset, defined as the proportion of agreement between a thematic map and reference data assumed to be correct "Ground Truth".

The diagonal of confusion matrix table lists the number of features that are classified into the correct ground truth class.

Three different metrics are usually calculated to assess the thematic accuracy (Congalton & Green, 1999):

(1) Overall accuracy: the ratio between the sum of the number of pixels classified correctly and the total number of pixels in the area of interest.

$$Overall\ accuracy = \frac{(i_{1,1} + i_{2,2} + i_{3,3} + i_{4,4})}{TOTAL}$$

(2) Producer's accuracy (P.A.):

$$Producer's\ accuracy = \frac{\text{Pixels correctly identified in reference given class}}{\text{Pixels correctly identified in reference class}}$$

(3) User's accuracy (U.A.):

$$User's\ accuracy = \frac{\text{Pixels correctly identified in a given map class}}{\text{Pixels claimed to be in that map class}}$$

Classification errors can be divided in two different categories:

$$Omission\ error = 100\% - P.A.$$

$$Comission\ error = 100\% - U.A.$$

The accuracy assessment method has been applied for each study of this work.

Chapter 3

Methodology

3.1 Surface Water Extraction

This section describes the employed surface water extraction method and the platform employed to build and deploy the algorithms. Firstly, Google Earth Engine and its capabilities and benefits will be described. Secondly, the section will detail the case studies over which the MWI is applied. Finally, after the validation phase the methodology will be applied to the entire Sentinel-2 catalog.

3.1.1 Google Earth Engine Capabilities and Application

The thresholding and application of the MWI algorithm as well as the selection of the correct images to process, has been done using Google Earth Engine.

Google Earth Engine is a planetary scale platform for Earth Science and data and analysis. The platform is entirely based on the cloud infrastructures and servers of Google. The platform combines a multi-petabyte catalog of satellite imagery and geospatial datasets with planetary-scale analysis capabilities and makes it available for scientists, researchers, and developers to detect changes, map trends, and qualify differences on the Earth's surface.

The platform harvests periodically the servers free satellite image providers like Landsat(USGS) and Sentinel(ESA). The archive includes more than thirty years of historical imagery and scientific datasets, that expands daily. Moreover, the data is instantly available.

The platform can be used online at the web page code.earthengine.google.com by creating a free account. The interface of the code editor is visible in **Figure 32**

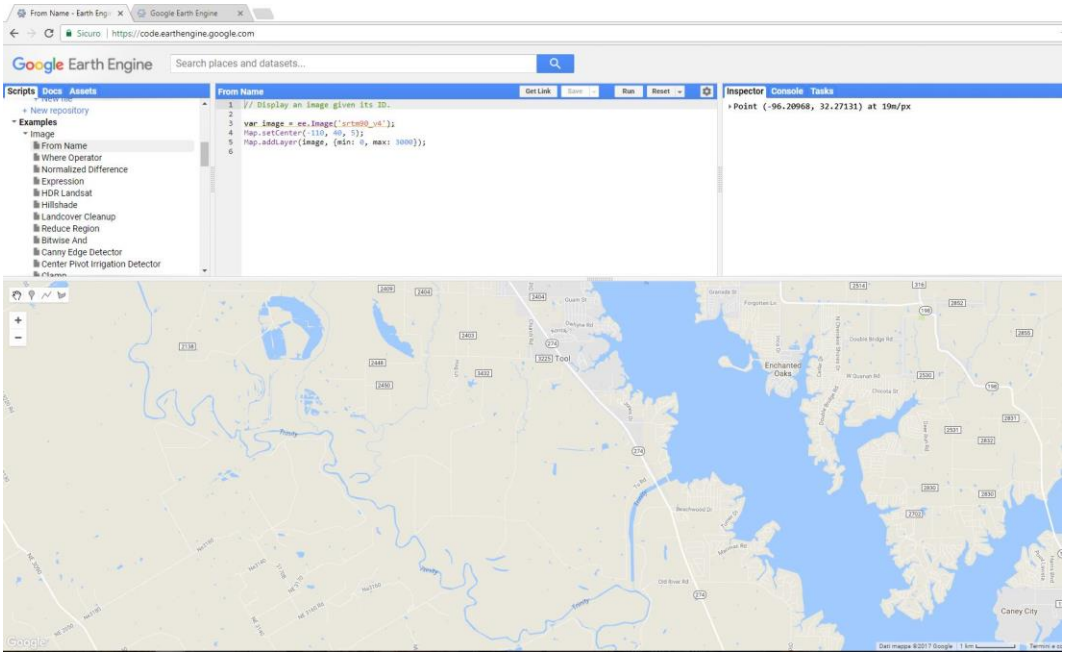


Figure 32 Earth Engine Interface

The platform permits to export the analysis in common formats like GeoTIFF and kmz that can be then integrated with other data sources or findings that are not available in Earth Engine.

3.1.2 Case Studies

The proposed procedure to extract surface water is based on extracting and validating the method on a single remote sensed image and later validating it using the extraction made from satellite images with higher GSD. The steps are summarized below:

1. Identification of the areas of interest over which is possible to extract the ground truth data using satellite images with resolutions that are higher or equal to the ones on which the water extraction methodology has been applied;
2. Extraction of the ground truth data that represents the water surface extent using means of visual interpretation;
3. Extraction of the water surface extent using the proposed methodology:
    a. Computation of NDWI, MNDWI and WRI over the area of interest;

b. Single threshold of NDWI, MNDWI, WRI, and sum of the three indexes which permits to obtain the MWI synthetic index that will be used for the surface water computation;
c. Selection of the DN equal to 3 that represent the areas where all 3 indexes extracted pixels classified as water;
d. Final surface water mask.

4. Quality assessment of the result by means of the error matrix computation.

At the current state of the art, five case studies were analyzed. As visible in **Figure 33** the areas of interest are located specifically in:

1. Sri Lanka – South East of Kalutara (Region of Kalutara) the Bentota Ganga River (area 1: Beruwala);
2. Italy – North East of Parma (Region of Lombardia) the Po river (area 2: Viadana);
3. Finland – North West of Kemi (Region of Lapland - Border between Finland and Sweden) the Torne river (area 3: Karunki);
4. Colombia – South West of Caucasia (Antioquia department) confluence between the Cauca river and Tarazà river (area 4: Puerto Antioquia);
5. Germany – South West of Plauen (Free State of Saxony) Droda and Pirk dam (area 5: Magwitz);

Figure 33 Location of the 5 case study area of interest

After the quality assessment of the water masks extracted over each area the computation is extended over all the collection of Sentinel-2 data in order to generate an image stack with all MWI analyzed bands.

3.1.2.1 Accuracy metrics and comparison to ground truth data

Table 2 summarizes the accuracy assessment done for each area using the different normalized indexes over each area of study. The complete accuracy assessment is visible in **Annex 1**.

As visible from **Table 2** the highest UA are provided by the MWI method described at ***2.3.2.1 Information Extraction Techniques***. It has been chosen to use this method because by providing high UAs permits to keep low the commission errors therefore we think it is better from an application point of view. For the area 4_Puerto Antioquia the UA is the lowest compared to the other areas. Nevertheless, the OA is almost the same as for the other areas. In this specific case the algorithm correctly classifies the pixels that are not water therefor we still have a high OA. One of the reasons could be the general characteristics of this specific S2 image with a higher quantity of clouds and waterbodies that have a higher quantity suspended matter.

	NDWI			MNDWI			WRI			thNDWI + thMNDWI			thNDWI + thMNDWI + thWRI		
	OA	PA	UA	OA	PA	UA	OA	PA	UA	OA	PA	UA	OA	PA	UA
1_Beruwala	99,25%	98,52%	95,53%	93,22%	98,56%	64,61%	99,54%	98,16%	98,09%	99,40%	97,32%	97,92%	99,44%	97,09%	98,28%
2_Viadana	64,58%	94,44%	12,34%	98,14%	77,96%	85,21%	98,45%	74,25%	95,09%	98,45%	76,70%	92,49%	98,44%	73,61%	95,52%
3_Karunki	98,92%	95,54%	95,10%	98,69%	92,96%	95,46%	98,60%	90,40%	97,20%	98,73%	92,37%	96,39%	98,60%	90,65%	97,14%
4_Puerto Antioquia	92,85%	94,55%	49,26%	96,92%	80,82%	76,27%	98,28%	85,19%	88,75%	97,48%	77,62%	84,77%	97,70%	75,41%	89,97%
5_Magwitz	88,34%	93,44%	16,33%	78,58%	95,85%	9,75%	99,52%	84,21%	95,07%	99,36%	92,83%	82,79%	99,52%	84,21%	95,07%

Table 2 Surface water extraction indexes accuracy

3.1.3 Data Analysis and Processing of Sentinel-2 data collections

The used datasets are the Sentinel-2: Multispectral Instrument (MSI), Level-1C with top of atmosphere reflectance scaled by 1000. Each scene has its granule identifier indicating its UTM grid reference. The metadata field is also populated including:

- CLOUDY_PIXEL_PERCENTAGE: granule-specific cloudy pixel percentage;
- CLOUD_COVERAGE_ASSESSMENT: cloudy pixel percentage for the whole archive that contains this granule.

The simple workflow is summarized below:

1. Bound filtering of all the acquired images on specific areas of interest;
2. Filtering of only the pixels that are cloud less;
3. Computation of the WRI, NDWI and MNDWI over each scene pre-selected at point 1 and 2 the 3;
4. Single threshold of NDWI, MNDWI, WRI, and sum of the three indexes which permits to obtain the MWI synthetic index that will be used for the surface water computation;
5. Thresholding of the MWI using the Otsu thresholding method;
    a. Optional for mountain areas: exclusion of commission errors using the height above nearest drainage dataset (Deltares);
6. Output display of the result.

3.1.3.1 Application of the methodology on the Sentinel-2 collection

1. Extraction of stable water surface, the areas where water never changes in all the scenes;
    a. Optional: Integration and comparison with existing datasets https://global-surface-water.appspot.com/;
2. Time series creation over river basins.

The S-2 Level-1C (TOA corrected) data is not furthermore processed and corrected atmospherically. As already mentioned in (Yun, et al., 2016) TOA better suits surface water extraction. TOA corrected data reduces the cosine effect due to different solar zenith angles and compensates the exomorphic solar irradiance (Wenbo, et al., 2013), (Byoung, Hyeong, & Jae, 2015).

The cloud mask with 60 m GSD has not been used because it does not suit the 10 m GSD proposed goal for this task.

The code implemented to validate and test the algorithm is attached as an annex at page 123.

3.1.4 Information Dissemination

At the current stage the most plausible option for the image dissemination is the Geonode Platform (http://geonode.org/). Although it's possible that I will build a platform that displays the results and then makes them easy to download in the most common formats, e.g. Esri shapefile and .kmz.

3.2 Building Footprint Extraction

The following section is based on the conference papers *Satellite Image Segmentation with Deep Residual Architectures for Time-Critical Applications* presented at the 26th European Signal Processing Conference of 2018 (EUSIPCO 2018) and *Building Footprint Extraction from VHR satellite imagery using a Deep Learning approach* presented at the IX Conference of the Italian Society of Remote Sensing (AIT 2018). Both works are done in collaboration with the Department of Electronics and Telecommunications of the Politecnico di Torino and the Joint Open Laboratories of Telecom Italia and the Politecnico di Torino. The goal of the work was to extract building footprints from VHR satellite imagery. The architecture of the algorithm based on Convolutional Neural Networks (CNN) is implemented by the Department of Electronics and Telecommunications. The creation of the datasets used for training and validation is done by the Author of the thesis (me).

3.2.1 Case Studies and datasets

In this section there will be described the datasets used for the work. The datasets used for the study are divided into two categories, training sets and test sets. The training sets are used to train the algorithm and the test sets are used to test the output of the algorithm against the ground truth data.

The test was carried out with a total of 9 VHR resolution images acquired by three different Earth Observation (EO) satellites. The images are acquired over 9 different areas worldwide. Each satellite image has four spectral bands (blue, green, red and IR) and a GSD of 0.5 m at nadir and different off-nadir angles. The reason we chosen the 9 datasets, 6 used as training set and 3 used as test sets, is that they are almost cloud free datasets.

The datasets and their description are listed below:

- 6 images used as training set:
    - D5 - Pléiades-1A © CNES (2014) GSD 0.5 m, 26.5° off-nadir angle (31122, 28969) n. of pixels (columns and rows);
    - D13 - Pléiades-1A © CNES (2015) GSD 0.5 m, 23° off-nadir angle (14923, 11445) n. of pixels n. of pixels (columns and rows);
    - D16 - Pleiades-1B © CNES (2017) GSD 0.5 m, 24.7° off-nadir angle n. of pixels (14538, 16382) (columns and rows);

    - D17 - Pleiades-1B © CNES (2016) GSD 0.5 m, 18° off-nadir angle n. of pixels (47725, 31475) (columns and rows);
    - D18 - Pleiades-1A © CNES (2016) GSD 0.5 m, 17.4° off-nadir angle n. of pixels (25843, 17530) (columns and rows);
    - D20 - WorldView-2 © DigitalGlobe (2016) GSD 0.5 m, 17° off-nadir angle n. of pixels (25843, 17530) (columns and rows);
- 3 images used as test set and validation:
    - D12 - Pleiades-1A © CNES (2016) GSD 0.5 m, 17° off-nadir angle original n. of pixels (26768, 21886) subset for validation (6301, 4855) (columns and rows);
    - D19 - WorldView-2 © DigitalGlobe (2012) GSD 0.5 m, 2.5° off-nadir angle original n. of pixels (9738, 9087) subset for validation (4683, 2423) (columns and rows);
    - D22 - Pleiades-1A © CNES GSD 0.5 m, 17.2° off-nadir angle original n. of pixels (29245, 19936) subset for validation (7601, 7301) (columns and rows);

In **Figure 34**, **Figure 35**, **Figure 36** show the three dataset used for the testing of the proposed CNN architecture. For validation purpose subsets of the areas that were highly covered by building footprints. Green areas, water bodies, shrubs and agricultural areas are also present in the subsets. Therefore, the subsets could be considered as appropriate to test the proposed method because it's heterogeneous.

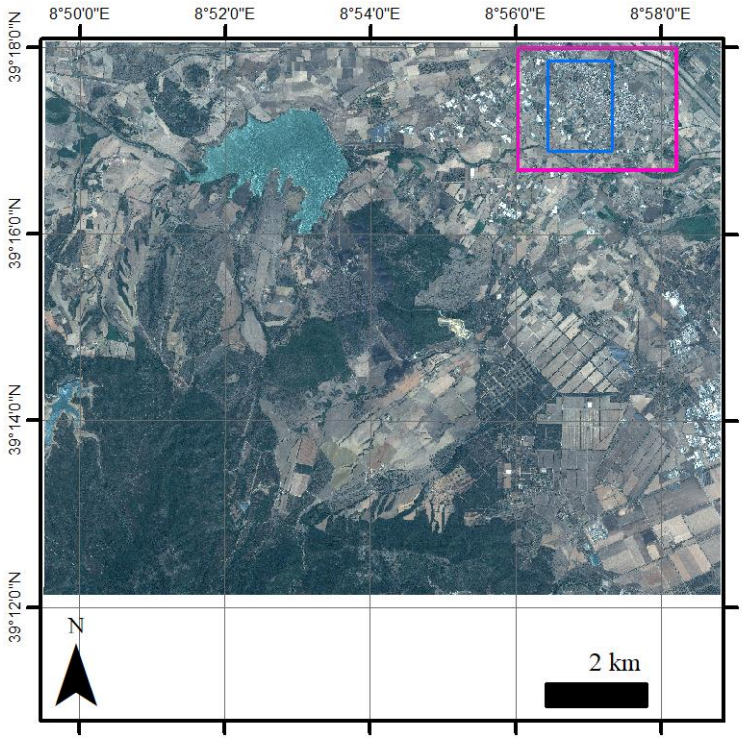


Figure 34 D12 test set image. in magenta the subset used for validation, in blue polygon the extract used for visualization purpose

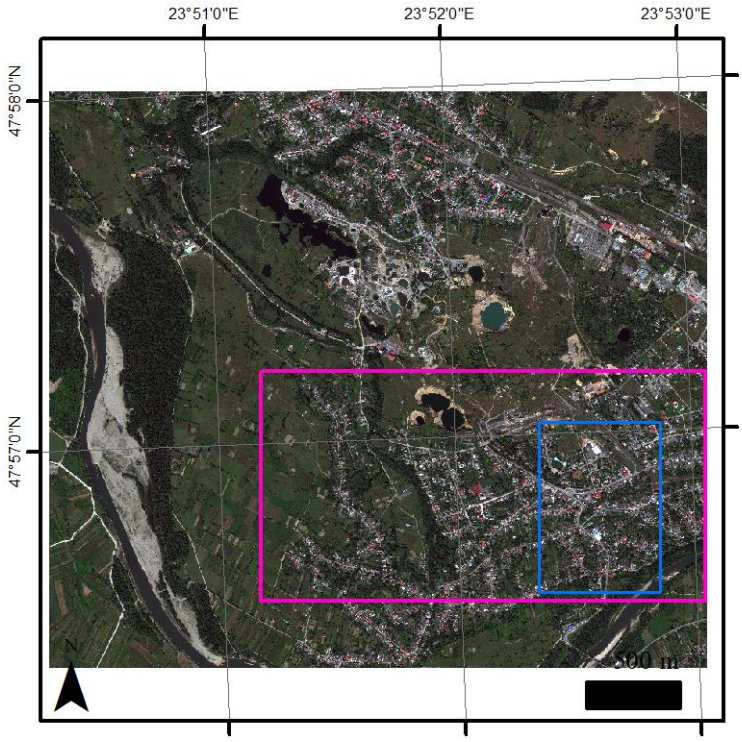


Figure 35 Detail of D19 test set image. polygon in magenta is the subset used for validation, blue polygon is the extract for visualization purpose

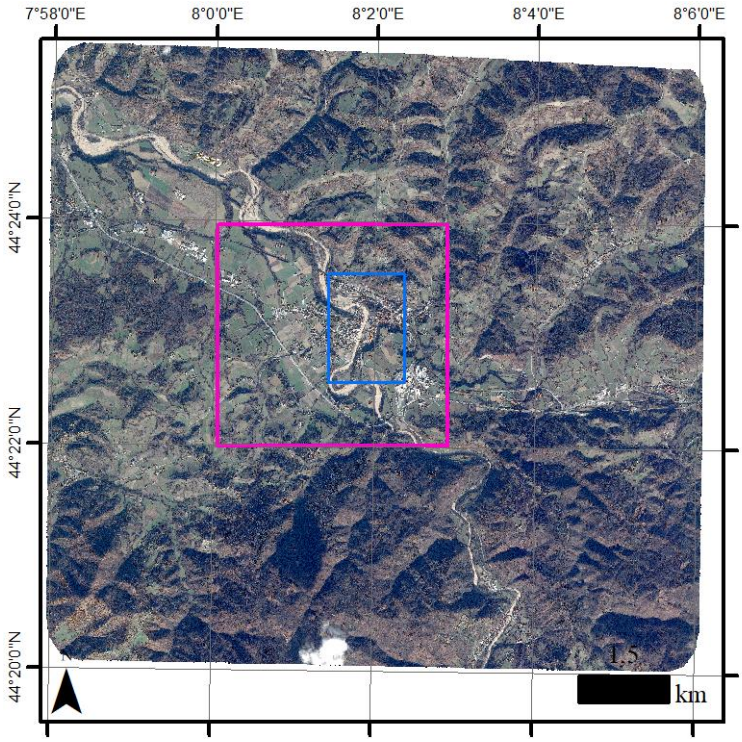


Figure 36 Detail of D22 test set image. polygon in magenta is the subset used for validation, blue polygon is the extract for visualization purpose

3.2.2 Data Analysis and Processing

3.2.2.1 Building footprint extraction methodology

The proposed network architecture (shown in **Figure 37**) consists of *encoder* and *decoder* modules which process the input image with a bottom-up (encoder) and top-down (decoder) approach. Firstly, encoder module extracts high level visual representations namely feature maps (feature maps are generate previously using the training samples generated from the 6 images used as training sets) from an input image. Secondly, decoder module takes as input the generated feature maps and predicts the corresponding segmentation score maps over several predefined classes. In the following the structure and function of each module is described in detail.

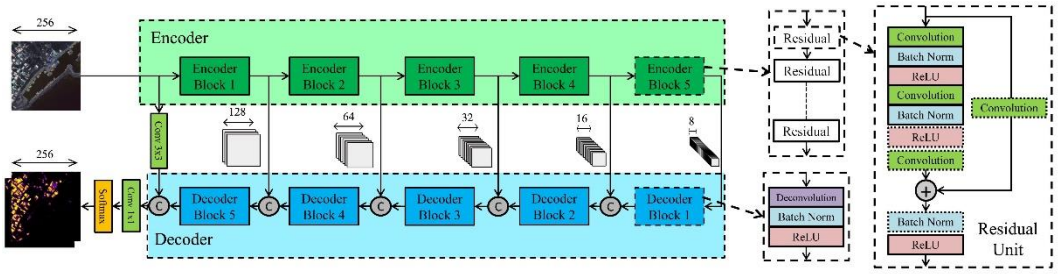


Figure 37 Proposed Network include encoder (top) and decoder (bottom) modules, an encoder and a decoder block are detailed in dashed boxes

Encoder: The proposed encoder module is arranged into five subsequent blocks and based on residual units. The residual encoder enables very deep architectures (up to 200 layers) while the training process remains feasible. We argue that deeper residual encoder enables learning high level visual features and results in better generalizations across samples with various semantic characteristics.

The encoder module takes as input an image sized 256 x 256 composed of 4 spectral channels and through five subsequent blocks the corresponding feature maps are extracted using residual units. The first layer of each block is a convolutional layer with stride of 2 pixels in which the generated feature maps are halved in spatial size while the number of feature maps increases with a specific factor depending on the encoder depth. The number of feature maps and convolutional layers for six different depths of encoder are provided in table 1. For instance, considering an encoder with depth of 18 layers, the first block takes as input 256 by 256 pixels input image, then the first block produces 64 feature maps of 128 by 128 pixels. Next, these feature maps are input to the subsequent block where 128 feature maps of 64 by 64 pixels are generated. This process continues at the next three blocks and in the end, the fifth block outputs 512 feature maps sized 8 by 8 pixels. That is, by proceeding to deeper layers in the encoder, higher level semantic patterns are extracted from the input and each block outputs increasing number of feature maps with decreasing resolutions. In addition, by proceeding to deeper blocks, the field of view of convolutional layers is expanding which results in learning visual representations in wider range.

Decoder: The decoder module is organized in five blocks symmetrical to the encoder and based on deconvolutional layers. Decoder takes as input the generated feature maps by encoder and predicts the segmentation score maps over land-use classes. Each decoder block is made up of a deconvolutional

ReLU and a batch normalization layer in which the feature maps are scaled up by factor of two while the feature maps number decreases. Let us consider the case of encoder with depth of 18 layers which detailed in previous section. The first decoder block takes as input 512 feature maps sized 8 by 8 pixels and generated by the last encoder block. Then, these feature maps are scaled up by factor of two while their number decreases resulting in 16 by 16 pixels of 256 feature maps. Next, through skip connection which are provided between each encoder block output and each decoder block input, these feature maps are concatenated with the feature maps with identical size generated by forth encoder block and then are input to second decoder block. These skip connections which provide the spatial information of higher resolution encoder feature maps in the decoder, are necessary to produce finer segmentation score maps. This process continues for the next three decoder blocks and the last block outputs 64 feature maps of 256 by 256 pixels matching the input size. The last convolutional layer is employed to perform a 1 by 1 convolution on the 64 feature maps and predicts the related confidence scores over each land-use class. Finally, since we are interested in class probabilities, a spatial SoftMax layer (Bridle, 1990) utilized to outputs the class score maps.

3.2.2.2 Network training

After defining the network architecture, we detail the process used for training the network. Firstly, the images in the dataset are divide into two sets: one set is used to generate training and validation samples while the other set is used for testing the trained network. Therefore, the train and test samples are extracted from different images indicating a statistically large difference between train and test samples. This difference stresses the proposed methodology and enables measuring the generalization capability of the network in a more practical condition in which the trained network expected to be applied over an image with no similarity with the training images. Next, to produce training samples, about 80% of each image in the first set is subdivided into 364 by 364 tiles by a translation of 120 pixels while the rest 20 % are used to extracted 256 by 256 validation samples with no translation. While the network input size is 256 by 256, the larger tiles used for training enables several label-preserving transformations during training such as random rotation and cropping. These transformations allow data augmentation improving the generalization capability of the network. Finally, the test tiles are extracted out of the second set of images and with the size of 512 by 512 which is the maximum size allowed by the GPU memory

(NVIDIA Titan X Pascal GPU 12 GB of RAM). Overlapping tiles are deployed to prevent the artifacts around tile boundaries.

Next, after generating the training samples, we proceed training the network by minimizing the following loss function:

$$L(w, y, t) = \sum_{i=1}^{HxW} \sum_{k=1}^{C} t_{i,k} \, log(y_{i,k})$$

where y and w are the network output and parameters respectively, t is the expected output and H, W and C are height, width and the number of classes respectively. In addition, to prevent overfitting and help generalization, we optimize the following function at training time:

$$J(w, y, t) = \eta L(w, y, t) + \lambda R(w)$$

where η and λ are learning rate and regularization parameters and R(w) is the L2 norm of all network parameters. The training is performed using Stochastic Gradient Descent with momentum of 0.9 and over mini-batches of 8 samples. The initial learning is 0.005 and divided by 5 every 30 epochs.

3.2.2.3 Training and testing sets

Training samples are extracted from the first set of images and covering 80 percent of each image while the validation samples are generated from the other 20 percent of each image. To measure the proposed network performance and to be able to compare it with existing classification algorithms available in commercial software, the test set is used.

3.2.2.4 Ground truth generation

For each image the ground truth data is generated by exploiting already existing building footprints downloadable from OpenStreetMap. The existing data is not always correct or complete and therefore for each dataset was necessary to manual refine the data wherever required. The ground truth is a binary mask which with pixel values equal to 1 represent building footprints and 0 other features. An example of the ground truth generation process can be visible in **Figure 38**, from left to right, in the first panel the original satellite image, central panel in violet the building footprints extracted from OpenStreetMap and in the final panel binary mask obtained by rasterizing the building footprints.

The ground truth is generated using the building footprints extracted from Open Street Map that are manually refined and integrated in case of missing data. Figure 2 provides 3 test images (D12, D19 and D22) each sized 6000 x 4500, 3700 x 2100 and 8700 x 6600 respectively

Figure 38 Ground truth extraction process

3.2.2.5 Accuracy metrics and comparison to existing procedures available in commercial software

The performance of the proposed CNN algorithm is compared to a supervised Mahalanobis distance classification which is performed over the three testing datasets. The Mahalanobis distance classification of each dataset is performed using only the region of interest (ROI) extracted from the dataset. ROI coming from other datasets are not used. For example, ROI extracted from D22 were used to perform Mahalanobis distance classification only on the D22 dataset. The chosen samples to perform the Mahalanobis distance algorithm represent: buildings with different type of roof colours (the number can vary between 2 and 4 depending on the analysed area), water bodies (rivers and lakes), paved roads, green areas and bare ground areas. The pixels that represent the buildings with different type of roofs are later added in one class that represent building footprints and attributed a pixel value equal to 1. Moreover, the pixels that represent the other classes are added together, to this result a pixel value equal to 0 is attributed. This procedure was applied for each dataset. The classification was performed in Envi 5.4.

3.3 Building damage assessment from vertical images

The following section is based on the paper Building damage assessment scale tailored to remote sensing vertical imagery published on the European Journal of Remote Sensing on 19th of October 2018. The study proposes a new building damage scale based on a quantitative thematic accuracy analysis of the data generated by the CEMS service after the Central Italy earthquake

of August 2016. The public available information extracted from VHR satellite data was analysed and compared to higher resolution ground truth data which permitted to propose a standard damage scale classification. The work demonstrates that by using different damage classes and detailed guidelines it is possible to obtain higher accuracies.

This paragraph will focus on the description of the case studies, the data processing and analysis and the methods of dissemination of the results.

3.3.1 Case Studies

On 24 August 2016 at 03:36:32 local time (01:36 UTC), an earthquake, measuring 6.2 on the moment magnitude scale, hit Central Italy in an area near the borders of the Umbria, Lazio, Abruzzo and Marche region. Its epicentre was close to Accumoli, with its hypocentre at a depth of 4 ± 1 km, approximately 45 km north of L'Aquila and 75 km southeast of Perugia (INGV, 2017). The official figures of the Protezione Civile report that the earthquake caused the death of 297 people: 11 in Accumoli, 49 in Arquata del Tronto and 234 in Amatrice. In addition to the loss of human lives, widespread destruction of cultural heritage was also reported.

3.3.2 Data Analysis and Processing

3.3.2.1 Identification of the satellite and aerial images used for damage assessment and data harvesting

As Copernicus EMS was the most active SEM mechanism during the Italian Earthquake events in 2016, the information and damage assessment generated in [EMSR177] Earthquake in Central Italy was analyzed in the first part. Initially the vector files of the satellite based damage assessment performed over AOI08 Accumoli [EMSR177] Accumoli: Grading Map, Monitoring 1 and AOI10 AmatriceWest [EMSR177] Amatrice West: Grading Map, Monitoring 1 were downloaded and the towns of Illica, Casale, Saletta, San Lorenzo e Flaviano and Accumoli (highlighted in red in **Figure 39**) were analyzed.

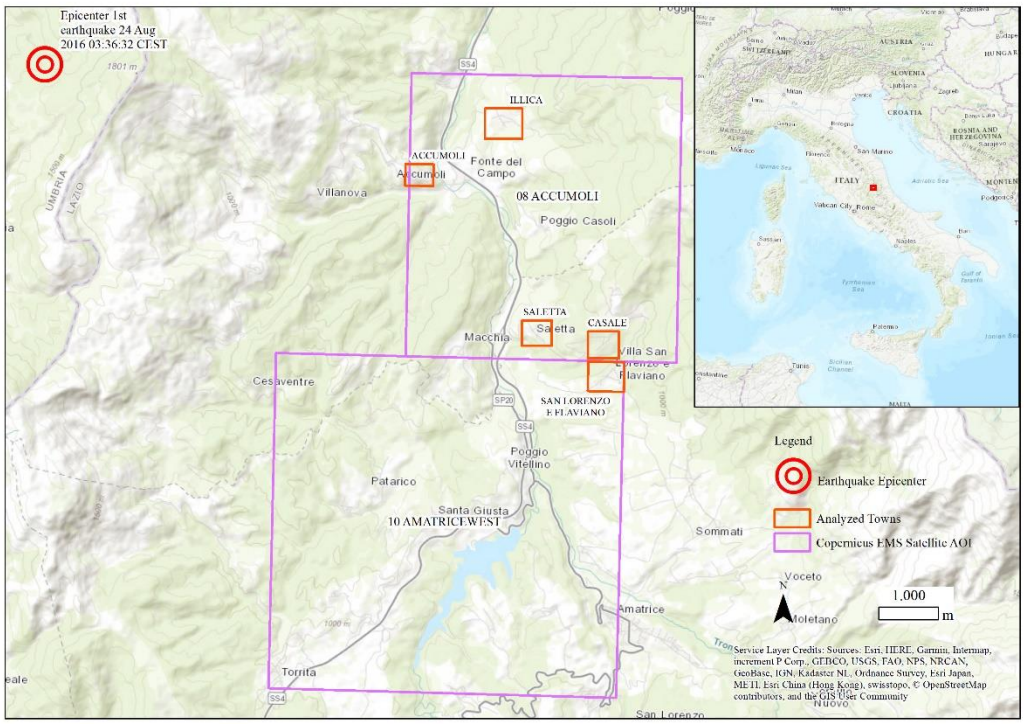


Figure 39 Analyzed areas

Copernicus EMS provides different types and formats of vector files related to reference and crisis data. From the vector zip package only the shapefile that represents the footprints of each building located in the areas of interest and including information on a satellite-based assessment of the building damages was analyzed. The damage scale uses five levels:

- Completely Destroyed;
- Highly Damaged;
- Moderately Damaged;
- Negligible to Slight Damage;
- Not Affected;

In **Figure 40**, as illustrative and informative purposes, the building footprints of Saletta's area of interest are displayed in a traffic light color code according to the damage scale class.

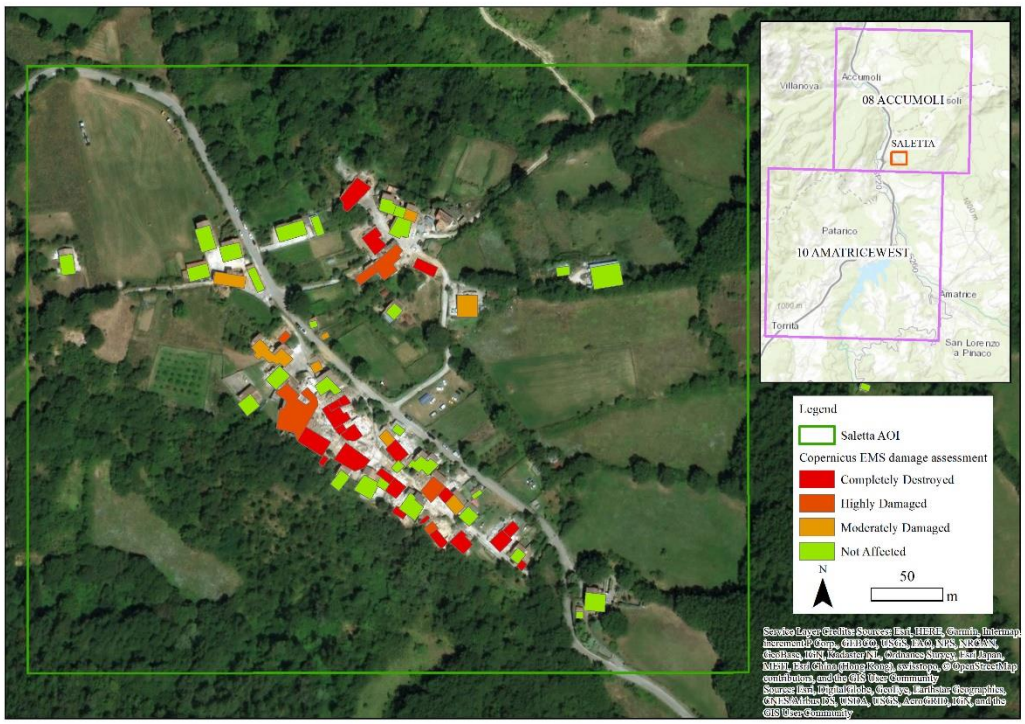


Figure 40 Damage detection of Saletta

From the vector metadata it emerged that Copernicus EMS's satellite damage assessment was based on the comparison between:

Pre-event image: Orthophoto 0.2 m © 2014 CONSORZIO TeA (formed by e-GEOS S.p.A. - CGR S.p.A. - Aerodata Italia srl).

Post-event image: WorldView-2 © Digitalglobe, Inc. (2016), (acquired on 25/08/2016 09:45 UTC, GSD 0.5 m, approx. 0 % cloud coverage, 34° off-nadir angle), provided under Copernicus by the European Union, ESA and European Space Imaging.

3.3.2.2 Ground Truth generation

Exploiting the availability of VHR post-event aerial imagery, a level of damage, using the same damage scale adopted by Copernicus EMS for EMSR177 was assigned to each single building and the related damage was described in an ad-hoc attribute field. Specifically, the analysis was carried out on the aerial images listed below:

- Pre-event image: Orthophoto 0.2 m © 2014 CONSORZIO TeA (formed by e-GEOS S.p.A. - CGR S.p.A. - Aerodata Italia srl).

- Post-event image: Aerial data © European Commission (acquired on 25/08/2016, GSD 0.1 m, 0 % cloud coverage) provided under Copernicus by CGR, Compagnia Generale Ripreseaeree (S.P.A.).

The visual image interpretation was based on the observation of: i) tone variations, which allowed to distinguish between different features; ii) shape variations, irregular building shape indicated structural building damage almost every time; iii) shadows, which may support the photo interpretation being a proxy of structure elevation. The result of this assessment was considered as Ground Truth for the analyzed case study, thanks to the higher GSD of the aerial image and the absence of time constraints for the analysis.

3.3.3 Information Dissemination

At the current stage the most plausible option for the image dissemination is the Geonode Platform. Although it's possible that I will build a platform that displays the results and then makes them easy to download in the most common formats, e.g. Esri shapefile and .kmz.

Chapter 4

Results

4.1 Surface Water Extraction

In this session it will be described the results obtained by the semi-automatic surface water extraction procedure over the 5 study areas. The outputs generated by algorithm will be analyzed qualitatively and quantitatively comparing it with the ground truth data.

Using the proposed method of surface water extraction over the area 1: Beruwala the OA is 99,44%, the PA and the UA are respectively 97,09% and 98,28%. In the **Annex 2** is visible how the algorithm classifies correctly the areas along the central river and in the area covered by sea water. The omission errors are mainly in the areas covered by the sea. The commission errors are visible along the coast and along the two branches of the main river. Although, these commission errors are debatable because as visible from **Annex 3** the data is classified correctly.

Over the area 2: Viadana the OA is 98,44%, the PA is 73,61% and the UA is 95,52%. As visible from **Annex 4** the method correctly classifies the main water body and the lakes in the east and west side of the image. The sandy areas are omitted by the classifier. This is due probably because the CEMS service classified the river and its bed as a water body **Annex 5**.

Over the area 3: Karunki the OA is 98,60%, the PA is 90,64% and the UA is 97,13%. As visible from **Annex 6** the proposed algorithm correctly classifies the main water body and the lake in the north-east part of the image. The omission errors are present along the riverside. The commission error is visible in the center of the river. But from **Annex 7** it is possible to see that there could have been possible interpretation errors from the CEMS service.

Over the area 4: Puerto Antioquia the OA is 97,70%, the PA is 75,40% and the UA is 89,97%. As visible from **Annex 8** the proposed methodology correctly extracts the water surface of the main river and its east tributaries.

Moreover, it also extracts the lakes in the south of the image. The omission errors are represented by the sandy areas of the riverbed and are visible over the east tributary river. Commission errors are visible in the center of the image, these areas have not been classified as water by the CEMS service. Although, it is visible from **Annex 9** that the areas have a chromatic response that looks like water.

Over the area 5: Magwitz the OA is 99,52%, the PA is 84,21% and the UA is 95,07%. As visible from **Annex 10** and **Annex 11** the Talsperre Pirk and the Talsperre Droda are correctly classified as water and the accuracy of the result is above 80%. The omission errors are visible along the banks of the reservoirs and for some small size waterbodies in the east part of the image. Commission errors are present in the south part of the image these areas have probably been classified as a quarry by the CEMS service. Therefore, are not considered as water in the downloaded vector products.

Finally, in addition to the results over each case study area, the procedure has been generalized and extended to extract the surface water over each Sentinel-2 acquisition over a given area of interest. For example, in the **Annex 12** it is visible the river evolution in an area with monsoon climate and the water frequency of each pixel. Moreover, by querying the pixels where with water low water frequency it is possible to obtain periods over which is better to focus the analysis.

4.2 Building footprint extraction

As described in (Ghassemi et al., 2018) it has been observed that depth of encoder plays a crucial rule in generalization of the network through learning visual representations of high semantic level which are not limited to training images. Authors showed that the network with the encoder including 152 layers obtains the best results. Moreover, this type of encoder outperforms existing neural networks like the standard U-Net with 19 and 35 layers. In this section, the results of the proposed method are compared to the outputs of a supervised Mahalanobis distance (MDist) classification based on region of interest (ROI) extracted from each training dataset. The classification was performed in Envi 5.4.

For the MDist classification the ROIs were manually extracted and used as training samples from each satellite image. The samples consisted in: i) footprints of buildings characterised by different type of roofs and tile colors (red, white and grey), ii) vegetation, iii) agricultural fields, iv) roads. An example of ROI is visible in **Figure 41**. After the classification each class has

been manually assigned to a binary class, Building/Other and later the accuracy assessment has been computed.

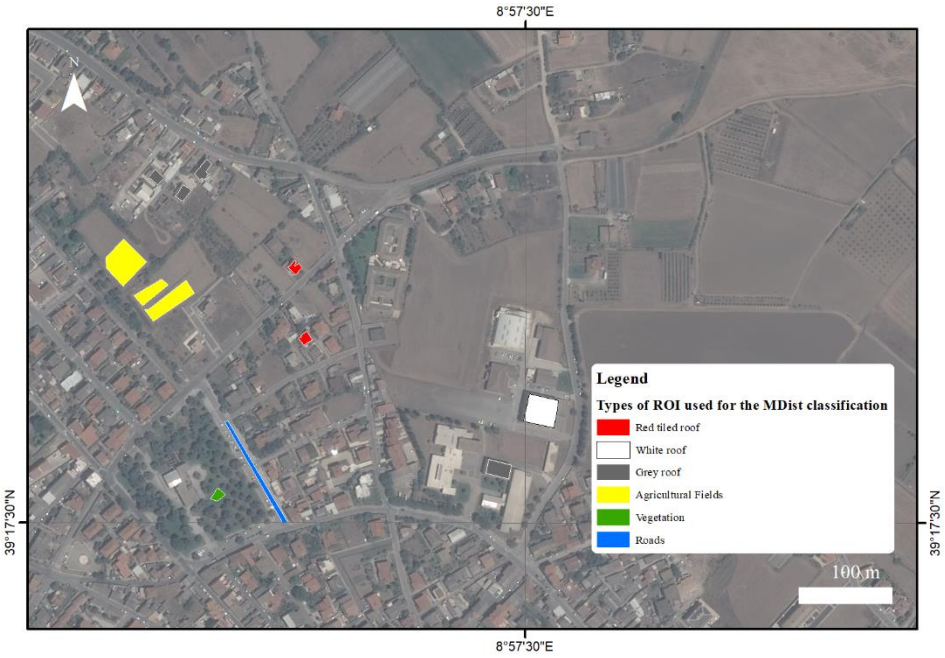


Figure 41 Types of ROI used for the MDist Classification (example from D12 dataset)

The classification was performed for each testing dataset and was then evaluated computing the respective Producer Accuracy (P.A.) and User Accuracy (U.A.) and compared to the P.A and U.A. of the results obtained with our proposed approach. The results are visible in

	PA CNN	PA MDist	UA CNN	UA MDist
D12	54,98	87,30	76,07	96,80
D19	58,90	52,36	71,90	61,80
D22	69,38	34,43	75,01	22,49

Table 3. Moreover, Omission Errors (O.E.) and Commission Errors (C.E.) have been also evaluated, as shown in **Table 4**. The results in Table 3 and 4 are related to a subset of the total building area of each image,

specifically: 69,8% for D12, 53% for D19 and 30,9% for D22 (blue polygon in **Figure 34 Figure 35 Figure 36**).

	PA CNN	PA MDist	UA CNN	UA MDist
D12	54,98	87,30	76,07	96,80
D19	58,90	52,36	71,90	61,80
D22	69,38	34,43	75,01	22,49

Table 3 Comparison of Producer and User Accuracy of the proposed method and a Mahalanobis distance supervised classification using Envi 5.4

	Omission CNN	Omission Mdist	Commission CNN	Commission Mdist
D12	45,02	12,70	23,93	3,20
D19	41,10	47,64	28,10	38,20
D22	30,62	65,57	24,99	77,51

Table 4 Comparison of the Omission and Commission errors of the proposed method and the Mahalanobis distance supervised classification using Envi 5.4

The results show that, apart from D12, our implementation outperforms the Mahalanobis distance algorithm this is more evident in the datasets D19 and D22. For D12 the user and the producer accuracy are noticeably higher, this is due to the characteristics of the satellite image. In fact, the image has a low off-nadir angle and the perimeter of the buildings is easier to distinguish by the Mahalanobis distance process. In

	PA CNN	PA MDist	UA CNN	UA MDist
D12	54,98	87,30	76,07	96,80
D19	58,90	52,36	71,90	61,80
D22	69,38	34,43	75,01	22,49

Table 3 it's visible how the proposed CNN architecture outperforms the Mahalanobis distance algorithm especially in the densely built area in the north-east of the image. In **Table 4** it's visible the commission and omission errors of each procedure. As per the PA and the UA Mahalanobis distance performs better only for D12. In **Figure 42** an extract of the MDist accuracy assessment over D22 it is visible that only four buildings are correctly classified. The MDist algorithm omits most of the building footprints and has commission errors along bright surfaces like roads. The proposed algorithm in comparison better extracts the building shapes, even if it tends to overestimate or underestimate along the building perimeters.

Concerning time constraints, the processing and the segmentation of the three test images took 110, 55 and 200 seconds for images D12, D19 and D22 respectively. The process is by far more fast compared to manually digitizing the images that usually takes about 10 hours each and 2 hours each for the Mahalanobis distance classification using a desktop pc with the following specifications (Intel(R) Core(TM) i7-7700 CPU @ 3.60GHz, 16GB of RAM, NVIDIA GeForce GT 710).

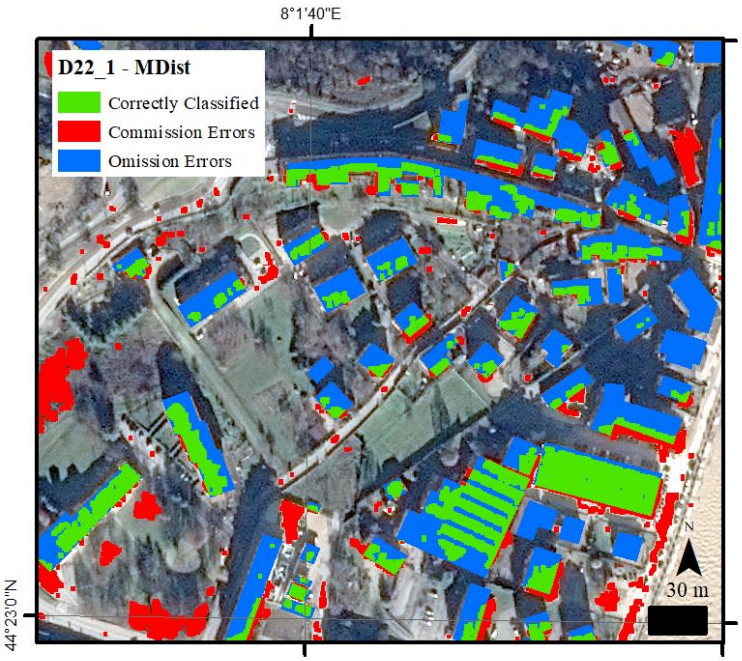


Figure 42 Results from the Mahalanobis distance algorithm of D22

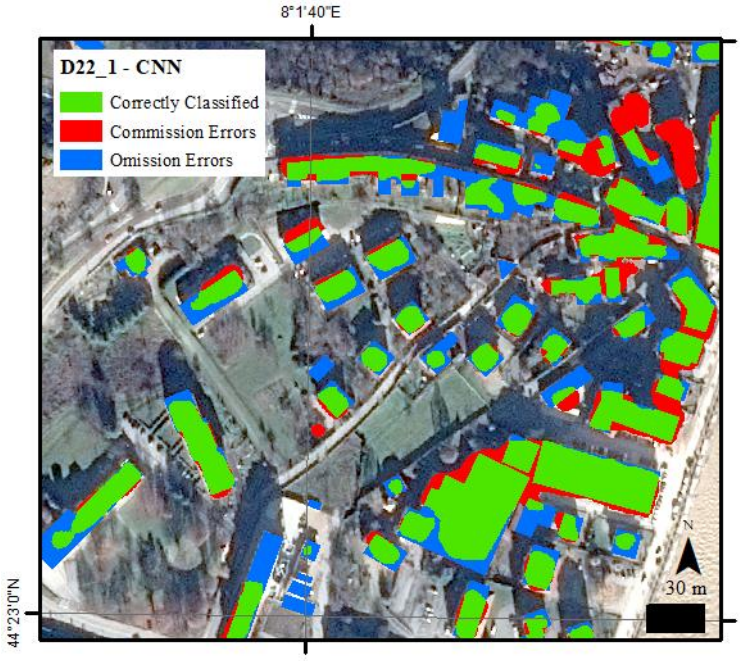


Figure 43 Results from the proposed CNN architecture of D22

4.3 Building damage assessment from vertical images - Operational Applications

The accuracy of Rapid Mapping analyses based on VHR satellite images carried out by Copernicus EMS was evaluated calculating the confusion matrices for each dataset/area. As already mentioned in the previous section (Independent aerial damage assessment), the results obtained by the Independent aerial damage assessment were considered as Ground Truth data.

The confusion matrixes and the related quality metrics related to the five examined areas are shown in **Figure 44**.

A) SALETTA		Copernicus Satellite Damage Assessment				
		Not Affected	Moderately Damaged	Highly Damaged	Completely Destroyed	Total
Independent Aerial Damage Assessment	Not Affected	17	4	0	0	21
	Moderately Damaged	8	1	2	0	11
	Highly Damaged	3	2	1	0	6
	Completely Destroyed	1	4	7	22	34
	Total	29	11	10	22	72

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	19%	41%	81%	59%
Moderately Damaged	91%	91%	9%	9%
Highly Damaged	83%	90%	17%	10%
Completely Destroyed	35%	0%	65%	100%

Overall Accuracy	57%

B) CASALE		Copernicus Satellite Damage Assessment				
		Not Affected	Moderately Damaged	Highly Damaged	Completely Destroyed	Total
Independent Aerial Damage Assessment	Not Affected	11	2	1	3	17
	Moderately Damaged	3	2	2	0	7
	Highly Damaged	0	3	1	2	6
	Completely Destroyed	3	1	2	19	25
	Total	17	8	6	24	55

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	35%	35%	65%	65%
Moderately Damaged	71%	75%	29%	25%
Highly Damaged	83%	83%	17%	17%
Completely Destroyed	24%	21%	76%	79%

Overall Accuracy	60%

C) ILLICA		Copernicus Satellite Damage Assessment				
		Not Affected	Moderately Damaged	Highly Damaged	Completely Destroyed	Total
Independent Aerial Damage Assessment	Not Affected	22	5	6	0	33
	Moderately Damaged	6	6	0	0	12
	Highly Damaged	6	12	2	11	31
	Completely Destroyed	1	6	2	27	36
	Total	35	29	10	38	112

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	33%	37%	67%	63%
Moderately Damaged	50%	79%	50%	21%
Highly Damaged	94%	80%	6%	20%
Completely Destroyed	25%	29%	75%	71%

Overall Accuracy	51%

D) ACCUMOLI		Copernicus Satellite Damage Assessment				
		Not Affected	Moderately Damaged	Highly Damaged	Completely Destroyed	Total
Independent Aerial Damage Assessment	Not Affected	114	0	0	0	114
	Moderately Damaged	23	0	0	0	23
	Highly Damaged	13	0	1	0	14
	Completely Destroyed	2	0	0	0	2
	Total	152	0	1	0	153

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	0%	25%	100%	75%
Moderately Damaged	100%	not applicable	0%	not applicable
Highly Damaged	93%	0%	7%	100%
Completely Destroyed	100%	not applicable	0%	not applicable

Overall Accuracy	75%

E) SAN LORENZO E FLAVIANO		Copernicus Satellite Damage Assessment					
		Not Affected	Negligible to slight damage	Moderately Damaged	Highly Damaged	Completely Destroyed	Total
Independent Aerial Damage Assessment	Not Affected	67	0	0	2	0	69
	Negligible to slight damage	0	0	0	0	0	0
	Moderately Damaged	18	0	2	1	0	21
	Highly Damaged	20	0	0	5	0	25
	Completely Destroyed	12	1	0	9	0	22
	Total	117	1	2	17	0	137

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	3%	43%	97%	57%
Negligible to slight damage	not applicable	100%	not applicable	0%
Moderately Damaged	90%	0%	10%	100%
Highly Damaged	80%	71%	20%	29%
Completely Destroyed	100%	not applicable	0%	not applicable

Overall Accuracy	54%

Figure 44 Confusion Matrixes related to the areas of Saletta(A), Casale(B), Illica(C), Accumoli(D), San Lorenzo e Flaviano(E). Copernicus EMS damage assessment vs Independent aerial damage assessment

It is observed that intermediate damage classes are those characterized by the largest discrepancies and lowest values, as confirmed by P.A. and U.A. in the confusion matrixes (A) (B) (C). P.A. values vary between a minimum of 6% (P.A. of Class Highly Damaged area C) and a maximum of 50% (P.A. of Class Moderately Damaged area C). U.A. values vary between a minimum of 9% (U.A. of Class Moderately Damaged area A) and a maximum of 25% (U.A. of Class Moderately Damaged area B).

In the matrixes (D) and (E) the P.A. and the U.A. do not follow the same patterns as per matrixes (A), (B), (C) due to: i) high percentage of not affected buildings; ii) high number of damaged buildings erroneously classified as not affected by the Copernicus EMS classification visible in **Figure 45**. The latter issue is also influenced by the intrinsic limitations of 0.5 m satellite imagery and the tight time constraint (few hours after post-event imagery availability) imposed by rapid mapping.

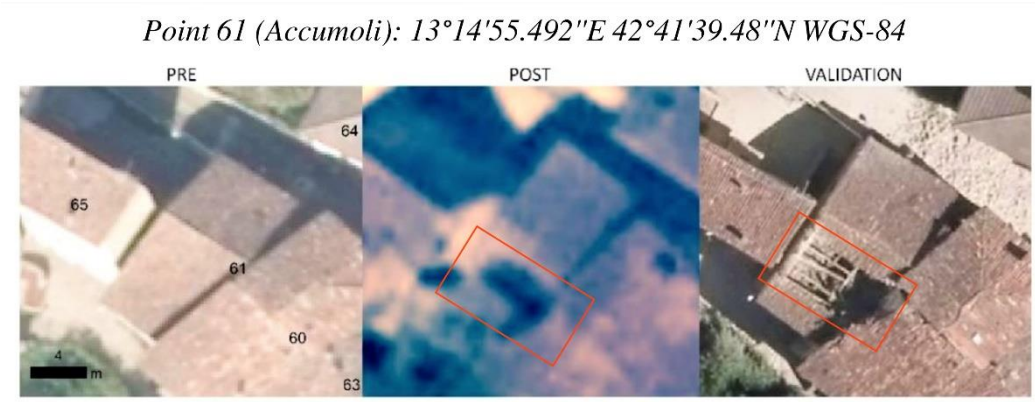


Figure 45 PRE: Accumoli pre aerial event image (GSD 0,2 m), POST: Accumoli post satellite event image (WV-2 sensor - GSD 0,5 m), VALIDATION: Accumoli post aerial event image (GSD 0,1 m)

The average overall accuracy is ~60%, value in line with the expected performance of photo interpretation of damages to buildings based on VHR satellite imagery (Corbane, Carrion, Lemoine, & Broglia, 2011).

4.3.1 Thematic accuracy evaluation

The obtained results permitted to identify the damage classes that lead to more discrepancies; specifically, the accuracy metrics of intermediate classes (Negligible to slight damage, Moderately Damaged and Highly Damaged) were significantly lower than in the other 2 classes (Not Affected and Completely Destroyed). Therefore, it was decided to perform a test by aggregating the aforementioned intermediate classes in a new class "Damaged" to evaluate the impact on accuracy metrics. The results are visible in **Figure 46**.

A) SALETTA		Copernicus Satellite Damage Assessment			
		Not Affected	Damaged	Completely Destroyed	**Total**
Independent Aerial Damage Assessment	Not Affected	17	4	0	21
	Damaged	11	6	0	17
	Completely Destroyed	1	11	22	34
	Total	29	21	22	72

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	19%	41%	81%	59%
Damaged	65%	71%	35%	29%
Completely Destroyed	35%	0%	65%	100%

Overall Accuracy	63%

B) CASALE		Copernicus Satellite Damage Assessment			
		Not Affected	Damaged	Completely Destroyed	**Total**
Independent Aerial Damage Assessment	Not Affected	11	3	3	17
	Damaged	3	8	2	13
	Completely Destroyed	3	3	19	25
	Total	17	14	24	55

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	35%	35%	65%	65%
Damaged	38%	43%	62%	57%
Completely Destroyed	24%	21%	76%	79%

Overall Accuracy	69%

C) ILLICA		Copernicus Satellite Damage Assessment			
		Not Affected	Damaged	Completely Destroyed	**Total**
Independent Aerial Damage Assessment	Not Affected	22	11	0	33
	Damaged	12	20	11	43
	Completely Destroyed	1	8	27	36
	Total	35	39	38	112

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	33%	37%	67%	63%
Moderately Damaged	53%	49%	47%	51%
Completely Destroyed	25%	29%	75%	71%

Overall Accuracy	62%

D) ACCUMOLI		Copernicus Satellite Damage Assessment			
		Not Affected	Damaged	Completely Destroyed	**Total**
Independent Aerial Damage Assessment	Not Affected	114	0	0	114
	Damaged	36	1	0	37
	Completely Destroyed	2	0	0	2
	Total	152	1	0	153

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	0%	25%	100%	75%
Damaged	97%	0%	3%	100%
Completely Destroyed	100%	not applicable	0%	not applicable

Overall Accuracy	75%

E) SAN LORENZO E FLAVIANO		Copernicus Satellite Damage Assessment			
		Not Affected	Damaged	Completely Destroyed	**Total**
Independent Aerial Damage Assessment	Not Affected	67	2	0	69
	Damaged	38	8	0	46
	Completely Destroyed	12	10	0	22
	Total	117	20	0	137

Class	Omission error	Commission error	P.A.	U.A.
Not Affected	3%	43%	97%	57%
Damaged	83%	60%	17%	40%
Completely Destroyed	100%	not applicable	0%	not applicable

Overall Accuracy	55%

Figure 46 Confusion Matrixes related to the areas of Saletta(A), Casale(B), Illica(C), Accumoli(D), San Lorenzo e Flaviano(E). Copernicus EMS damage assessment ((aggregation of classes Negligible to slight damage, Moderately Damaged and Highly Damaged to “Damage”

As expected P.A. and U.A. values of the aggregated Damaged class have increased, specifically:

P.A. is in the range from 35% (P.A. area A) to 62% (P.A. area B)

U.A. is in the range from 29% (U.A. area A) to 57% (U.A. area B)

Furthermore, a general increase of the overall accuracy is observed after the aggregation, i.e. from 57% to 63% (area A), from 60% to 69% (area B) and from 51% to 62% (area C). The outcomes confirm that intermediate classes are the ones affected by higher interpretation uncertainty.

Contrariwise, the P.A., U.A. and overall accuracy of the areas of Accumoli (D) and San Lorenzo e Flaviano (E) are not influenced by the aggregation, due to the aforementioned Copernicus EMS classification issues.

The results confirm the need to redefine also the damage interpretation guidelines, shifting from a structural damage (EMS-‘98 like) to damage (and damage proxies) visible on VHR vertical imagery.

Imagery spatial resolution and viewing angle are the two key factors to be considered in developing the damage scale proposal; in fact vertical imagery, i.e. almost null off-nadir angles, does not allow the facades of the buildings to be analysed or cracks and failure in walls to be detected.

4.3.2 Proposal of standard building damage scale tailored to remote sensing vertical imagery

According to the outcomes of the thematic accuracy evaluation, a new standard damage scale to be used for building damage assessment tailored to remote sensing vertical imagery is proposed in **Annex 13**, defining 4 damage classes.

Destroyed: assigned to structures that are totally or largely collapsed (>50%). This category shall be assigned also when only a portion of the building has collapsed to the ground floor. In these cases, the original building structure is no longer distinguishable.

Damaged: it shall be used when post satellite imagery is available and includes:

Major visible damages, which shall be assigned to structures with part of the roof collapsed and serious failure of walls;

Minor visible damage level, i.e. buildings with a largely intact roof characterized by presence of partial damage (collapse of chimneys or detach of roof tiles) or surrounded by large debris/rubble or sand deposit.

The separation between Minor and Major Damage grades can be used only when imagery with a GSD of approximately 0.1 m is available (typical for aerial and UAV imagery).

Possibly Damaged: it shall be used for buildings whose interpretation is uncertain, due to lower image quality (e.g. shadow or degraded resolution due to high off-nadir angle) or to the presence of possible damage proxies like small traces of debris/rubble or sand deposit around the building. This class attribution can be given by inferring the state of the building from surrounding features. In flooding it could be traces of water currents leading up to and then leaving a building or set of buildings.

No Visible Damage: it shall be assigned to the structures that appear to have complete structural integrity, i.e. when the walls remain standing and the roof is virtually undamaged. It is important to remark that this class don't exclude the presence of structural damages, i.e. the building may anyway have suffered damages that cannot be assessed from vertical satellite imagery regardless of its spatial resolution.

4.3.3 Application of the new damage scale on a different case study

The proposed building damage scale was validated. Afterwards it was decided to apply the new building damage scale to assess the earthquake's impact on Pescara del Tronto, another town affected by the seismic event in Central Italy.

4.3.4 Ground Truth generation: Independent UAV damage assessment

The Ground Truth data was identified by performing a UAV-based damage assessment based on the comparison between the images listed below.

Pre-event image: Orthophoto 0.2 m © 2014 CONSORZIO TeA (formed by e-GEOS S.p.A. - CGR S.p.A. - Aerodata Italia srl).

Post-event image: UAV 0.7 m Team Direct (Disaster RECovery Team) of Politecnico of Turin (acquired and processed on 07/09/2016).

The UAV damage assessment revealed that 18 structures showed no visible damage, 6 structures were classified as Possibly Damaged, 67 structures were damaged (40 structures suffered minor visible damage and 27 structures major visible damage) and 123 structures were destroyed. The output map is visible in **Annex 14**.

4.3.5 Damage class aggregation of Copernicus EMS vs Ground Truth

The Copernicus EMS damage assessment was downloaded from the [EMSR177] Grisciano Grading Map, Monitoring 1. The downloaded datasets were based on the analysis of the images listed below.

Pre-event image: Orthophoto 0.2 m © 2014 CONSORZIO TeA (formed by e-GEOS S.p.A. - CGR S.p.A. - Aerodata Italia srl).

Post-event image: WorldView-2 © DigitalGlobe, Inc. (2016), (acquired on 25/08/2016 09:45 UTC, GSD 0.5 m, approx. 0 % cloud coverage, 34° off-nadir angle), provided under Copernicus by the European Union, ESA and European Space Imaging.

Preliminary, the damage classes were aggregated in order to enable the comparison of the Copernicus EMS results with the Ground Truth. The adopted class aggregation is visible in **Figure 47**.

PESCARA DEL TRONTO		Copernicus Satellite Damage Assessment				
		No visible damage	Possibly Damaged	Damaged	Destroyed	**Total**
Independent UAV Damage Assessment	No visible damage	15	0	3	0	18
	Possibly Damaged	3	0	1	2	6
	Damaged	45	0	10	12	67
	Destroyed	10	0	8	103	121
	Total	73	0	22	117	212

Class	Omission error	Commission error	P.A.	U.A.
No visible damage	17%	79%	83%	21%
Possibly Damaged	100%	0%	0%	0%
Damaged	85%	55%	15%	45%
Destroyed	15%	12%	85%	88%

Overall Accuracy	60%

Figure 47 Copernicus Damage assessment compared to the ground truth

The accuracy assessment was performed using a confusion matrix as per the previous analysis.

Like in the previous analysis conducted by Copernicus EMS, the overall accuracy is about 60%, even aggregating the critical damage classes. This result is due to the interpretation uncertainty of the intermediate classes as confirmed by the low P.A. (Possibly Damaged 0%, Damaged 15%) and U.A. values (Possibly Damaged 0%, Damaged 45%) of **Figure 47**. This result confirm once again that also different interpretation guidelines should be developed and adopted.

4.3.6 Independent classification vs Ground Truth

The last validation is based on the generation of a new satellite-based damage assessment adopting the Proposed Building Damage Scale and the related interpretation guidelines (detailed in **Annex 13**) without time constraints. The images for the analysis are the same used by Copernicus EMS and listed in the previous section. The results were therefore compared to the Ground Truth generated from the UAV analysis. The accuracy assessment results are summarised in the Confusion Matrix shown in **Figure 48**.

PESCARA DEL TRONTO		Independent Satellite Damage Assessment				
		No visible damage	Possibly Damaged	Damaged	Destroyed	**Total**
Independent UAV Damage Assessment	No visible damage	18	0	0	0	18
	Possibly Damaged	1	5	0	0	6
	Damaged	24	12	31	0	67
	Destroyed	0	1	2	120	123
	Total	43	18	33	120	214

Class	Omission error	Commission error	P.A.	U.A.
No visible damage	0%	58%	100%	42%
Possibly Damaged	17%	72%	83%	28%
Damaged	54%	6%	46%	94%
Destroyed	2%	0%	98%	100%

Overall Accuracy	81%

Figure 48 Damage assessment using the proposed Building Damage Scale

The P.A. reports high values for the classes No visible damage (100%), Possibly Damaged (83%) and Destroyed (98%) and a 46% for the Damaged class. The U.A. reports high values for the classes Damaged (94%) and Destroyed (100%). The No visible damage class has 42% accuracy and the Possibly Damaged class reports 28%. As expected the accuracy is lowest in the Possibly Damaged class. This is plausible, and depending on the characteristics of the satellite image, principally the spatial resolution (lower than the UAV spatial resolution used to extract the Ground Truth) and the atmospheric conditions (haze) which mostly affected the CAPI.

Nevertheless, the overall accuracy is 81%, i.e. an increase of about 20% with respect to the Copernicus EMS damage assessment.

Chapter 5

Application Context

5.1 Surface Water Extraction

Surface freshwater available in lakes and rivers is a fundamental resource and plays an important role for the wellbeing and survival of life on Earth. Rivers represent 0.49% of the total surface freshwater and are one of the most important sources of water for humans. The proposed application tries to extract and represent rivers and lakes, acknowledging their importance and their fundamental role as a resource.

The proposed application can detect and monitor changes of waterbodies over the lifespan of 7.25 years of the Sentinel-2 satellites (from 23rd of June 2015 for Sentinel-2A and the 7th of March 2017 for Sentinel-2B). Potentially the application can provide data up until June 2024.

The water surface areas can be integrated within the existing Copernicus Global Land Service products as additional information that complements the water level measurements. The water level measurements that are provided by the Sentinel-3A satellites which have been recently harmonized and referenced to the geoid EGM 2008 the same used by the Jason 2 and 3 missions.

Update existing databases can be another field of application. For example, in Regione Piemonte the surface water cartographic elements are derived from maps that have been generated between 1991 and 1995. Therefore, an update surface waterbody map can be useful for different areas of application. An updated geodatabase can provide useful information for institutions, agricultural firms and can be generally used for territorial planning and monitoring. This approach can also be used to create new databases over areas and parts of the world where there are no surface water vector files available. For example, it can be used to update OpenStreetMap.

In the context of emergency management having an up to date surface water dataset can be useful to define the permanent water bodies. After a

flooding event the permanent water body can be subtracted from the standing water in order to define the flooded areas. This can be useful by using the proposed method and create on demand permanent water datasets for emergency management.

5.2 Building Footprint Extraction

Building footprint extraction procedure is described in the section 3.2 Building Footprint Extraction. Building footprints are usually employed in different type of analysis. Up to date building footprints are important in case of mapping remote areas, monitoring the construction of new buildings and are also useful in case of emergency management.

Mapping building footprints using satellite images is an activity that is time consuming and could be performed using the proposed algorithm. The application of the method can be done over built up areas of cities that are expanding rapidly.

Monitoring sensible areas and abusive constructions is another important task that can be carried out using the proposed method. Soil consumption represents an important issue for national and regional authorities, therefore a monitoring tool based at the single building level can be useful in case of land monitoring.

During rapid mapping of damages is crucial to have a up to date prior event building footprints. In this instance single building digitalization is requires a lot of effort and is one of the most time-consuming activities. The extracted building footprints are afterwards compared to the post-event image and are later classified according to existing damage scales taking in consideration the damage that have undergone.

5.3 Building damage assessment from vertical images - Operational Application

The proposed standard building damage scale has been accepted by the Copernicus Rapid Mapping Emergency Management Service and by the International Working Group of Satellite Emergency Mapping.

To date (8/11/2018) the standard scale has been applied by the Copernicus Rapid Mapping Emergency Management Service for the following activations and building damage grading maps:

- [EMSR257] Flood in Attika, Greece (November 2017);
- [EMSR260] Flood in Northern Italy (December 2017);
- [EMSR269] Tropical Cyclone GITA in Tonga (February 2018);
- [EMSR304] Earthquake in Lombok, Indonesia (August 2018);
- [EMSR317] Earthquake in Indonesia (September 2018);
- [EMSR320] Earthquake in Northwestern coast of Haiti (October 2018);

An extract of the maps is visible in the annexes **Annex 18**, **Annex 19**, **Annex 20**, **Annex 21**, **Annex 22** and **Annex 23**

In the specific case of the CEMS the performance of the damage scale has not yet been validated on each activation. Although, possible future goals could be the validation of the generated outputs in comparison to ground data or higher resolution images.

The proposed building damage scale could also be used as guidelines for VGI mapping projects. VGI based projects like HOT and Tomnod could benefit from the use of a standard damage grading scale. The use of the scale could possibly permit a more homogenous classification of the affected buildings after an earthquake or a catastrophic event.

The building damage scale could be employed for the validation of semi-automatic change detection algorithms.

Conclusions and further development

This thesis presents 3 different works related to the use and employment of remote sensed data. In specific, the three studies are related to the use of remote sensed data acquired from passive sensors with different technical characteristics: GSD, radiometric resolution, temporal resolution, etc. The different datasets have been described with the purpose to be adopted for emergency management applications and environmental assessment. The 3 works explore the usage of satellite data for application from the small to the large map scale. The studies can be summarized below:

1) Surface water extraction using medium GSD data: this work has been carried out using already existing surface water detection procedures, e.g. NDWI, MNDWI and WRI applied to Sentinel-2 images in a cloud computing environment (Google Earth Engine). To achieve the results, it has been proposed a new surface water extraction index derived from the combination of the three existing indexes mentioned before. Furthermore, the Otsu thresholding method is used to automatically differentiate the pixels that represent water from those that are not water;
2) Building footprint feature extraction: the work aims to extract building footprints using deep neural network algorithms. The procedure has been applied on VHR images with a GSD of 0.5 m and off-nadir angles that range from a minimum of 2.5° to a maximum of 26.5°. During the process the algorithm has been trained on 6 VHR satellite images and deployed on 3 VHR satellite images different from the ones used as the training set;
3) Building damage assessment from vertical images: with this study it has been proposed a standard building damage scale intended to be used over VHR satellite, aerial and UAV images. The work addresses the need of a standard building damage scale that permits to better

differentiate between damaged and not damaged structures after the catastrophic events (earthquakes, tornados, flash floods).

Each of the three studies has been evaluated from a performance point of view. The key results and the possible implications can be summarized as follows:

1) Surface water extraction using medium GSD data: on single satellite images the propose algorithm has a maximum OA of 99,52% and a minimum OA 97,70%. The maximum of PA is 97,09% and the minimum PA is 73,61%. And the maximum UA is 98,28% and the minimum UA is 89,97%. The evidence may imply that the method can be applied over the 5 analyzed case study areas and extended worldwide after a more detailed accuracy assessment;
2) Building footprint feature extraction: the analyzed method compared to single extraction using commercial algorithms from existing software shows that the proposed method outperforms the Mahalanobis in 2 out of the 3 tests done. The proposed procedure is promising but must be further developed and tested in other case study areas;
3) Building damage assessment from vertical images: the proposed standard building damage assessment scale performs about 20% better compared to the previous existing damage scale used by the CEMS service during RM activities.

The proposed work about water surface extraction joins existing works and aims to provide an integration or a follow-up to existing studies (e.g. (Pekel, Cottam, Gorelick, & S. Belward, 2016), (Donchyts, van de Giesen, & Gorelick, 2017)).

Building footprint extraction aims to propose a follow-up and an integration with the methods proposed by Microsoft (Bing, 2018) with an extend of the work to other areas of the world.

Building damage grading standard scale fits within the purpose of rapid mapping activities and thanks to its adoption as a standard by the IWG-SEM could be used by the voluntary mapping services and by the other rapid mapping map providers.

The proposed work tries to provide a contribution about surface water evolution and it's changes during time. Moreover, it provides operational tools for rapid mapping and for scholars that are interested in water bodies morphology and its evolution. The building footprint extraction procedure is a promising application and should be refined and extended using more case studies with more VHR acquisitions. Finally, the building damage scale applied to vertical imagery after it became operational has been used in different instances by the CEMS RM service.

Although this work proposed implementations to fields of work related to environmental assessment and rapid mapping there still are some gaps and limitations.

The Surface water extraction work is currently carried out using Google Earth Engine. Although, the platform provides immense computation power compared to personal work stations the platform is managed by a private company, therefore it could change terms of reference, could change substantially the code libraries or could be closed without notice. It could result in the loss of all the personal algorithms and procedures available on the servers. This could mean that the user should need migrate or revisit conspicuous parts of his/her algorithms. A remedy could be the use of pay per use platforms e.g. Sentinel Hub EO-Browser or ONDA Data Information Access Services (DIAS). Moreover, at the time of writing 7th of November 2018 Sentinel-2 data is not yet meeting the maximum 3 m multi-temporal registration error (confidence level of 95.5%) (ESA, 2018). Therefore Sentinel-2 is not meeting yet it's contract standards. Moreover, Sentinel-2 is a passive sensor, therefore the acquired scenes might be covered by clouds. This could compromise a possible monitoring service over lakes or waterbodies. Future studies should explore the use of Sentinel-1 data for the surface water extraction and the use of Sentinel-2 data for validation purpose.

The building footprint extraction should be further teste and future work needs to be done on other geographic areas and compared it to other algorithms available in commercial software. Moreover, it's crucial to find a vectorization method that suits our needs. This will permit to evaluate in a more robust way the proposed work.

The proposed building damage scale from vertical imagery has become a standard and accepted by the IWG-SEM. Future work should focus on the integration of the scale with automatic algorithms and on the application of

the scale to imagery with of nadir angles that permit to view the facades of buildings.

References

© European Union. (2012 - 2018, January 01). *Copernicus*. Tratto il giorno December 04, 2018 da Copernicus Emergency Management Service - Mapping: https://emergency.copernicus.eu/mapping/ems/ems-rapid-mapping-products

Ajmar, A., Boccardo, P., Disabato, F., & Giulio Tonolo, F. (2015). Rapid Mapping: geomatics role and research opportunities. *Rendiconti Lincei*(26), 63-73.

Ali, A. L., & Schmid, F. (2014). Data Quality Assurance for Volunteered Geographic Information. In M. Duckham, E. Pebesma, K. Stewart, & A. Frank, *Geographic Information Science. GIScience 2014. Lecture Notes in Computer Science* (pp. 126-141). Cham: Springer.

Ali, A. L., Schmid, F., Al-Salman, R., & Kauppinen, T. (2014). Ambiguity and plausibility: managing classification quality in volunteered geographic information. *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (pp. 143-152). New York, USA: 2014.

Angiuli, E., & Trianni, G. (2014). Urban Mapping in Landsat Images Based on Normalized Difference Spectral Vector. *IEEE Geoscience and Remote Sensing Letters, 11*(3), 661-665.

Attarzadeh, R., & Momeni, M. (2012). Object-Based Building Extraction from High Resolution Satellite Imagery. *XXII ISPRS Congress.* Melbourne, Australia: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences.

Awad, M., Chehdi, K., & Nasri, A. (2007). Multicomponent Image Segmentation Using a Genetic Algorithm and Artificial Neural

Network. *IEEE Geoscience and Remote Sensing Letters, 4*(4), 571-575.

Barron, C., Neis, P., & Zipf, A. (2014). A comprehensive framework for intrinsic OpenStreetMap quality analysis. *Transactions in GIS, 18*, 877-895.

Bhatti, S. S., & Tripathi, N. K. (2014). Built-up area extraction using Landsat 8 OLI imagery. *GIScience & Remote Sensing, 51*(4), 445–467.

Bing. (2018, June 28). *Bing Blogs*. Tratto il giorno December 05, 2018 da Bing Blogs: https://blogs.bing.com/maps/2018-06/microsoft-releases-125-million-building-footprints-in-the-us-as-open-data

Birkmann, J., & Welle, T. (2016). The WorldRiskIndex 2016: Reveals the Necessity for Regional Cooperation in Vulnerability Reduction. *Journal of Extreme Events, 3*(1). Tratto da http://www.uni-stuttgart.de/ireus/dateiuploads/Birkmann-and-Welle_2016_WorldRiskIndex2016.pdf

Bridle, J. (1990). Probabilistic Interpretation of Feedforward Classification Network Outputs, with Relationships to Statistical Pattern Recognition. In F. Soulié, & J. Hérault, *Neurocomputing. NATO ASI Series (Series F: Computer and Systems Sciences* (Vol. 68, p. 227-236). Berlin: Springer.

Byoung, C. K., Hyeong, H. K., & Jae, Y. N. (2015). Classification of Potential Water Bodies Using Landsat 8 OLI and a Combination of Two Boosted Random Forest Classifiers. *Sensors*, 13763-13777. doi:10.3390/s150613763

Carmen, B., & Bucher, B. (2010). Quality in user generated spatial content: A matter of specifications. *13th AGILE International Conference on Geographic Information Science 2010* (pp. 1-8). Guimarães, Portugal: AGILE.

Centre for Research on the Epidemiology of Disasters(CRED). (2015, January 01). *EM-DAT The International Disaster Database.* Retrieved September 08, 2017, from emdat.be: http://emdat.be/human_cost_natdis

Chatterjee, C., Kumar, R., Chakravorty, B., Lohani, A., & Kumar, S. (2005). Integrating Remote Sensing and GIS Techniques with Groundwater Flow Modeling for Assessment of Waterlogged Areas. *Water Resources Management, 19*(5), 539–554. doi:https://doi.org/10.1007/s11269-005-2071-4

Congalton, R. G., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: Principles and practices.* Boca Raton: Lewis Publishers.

Copernicus Europe's eyes on Earth. (2017, September 08). *Copernicus*. Tratto il giorno September 08, 2017 da Copernicus Global Land Service: http://land.copernicus.eu/global/about

Corbane, C., Carrion, D., Lemoine, G., & Broglia, M. (2011). Comparison of Damage Assessment Maps Derived from Very High Spatial Resolution Satellite and Aerial Imagery Produced for the Haiti 2010 Earthquake. *Earthquake Spectra, Vol. 27*(No. S1), S199-S218.

Donchyts, G., van de Giesen, N., & Gorelick, N. (2017). Reconstruction of reservoir and lake surface area dynamics from optical and SAR satellite imagery. *9th International Workshop on the Analysis of Multitemporal Remote Sensing Images (MultiTemp).* Bruges, Belgium. Tratto da https://www.researchgate.net/publication/318128244_Reconstruction_of_reservoir_and_lake_surface_area_dynamics_from_optical_and_SAR_satellite_imagery

ESA. (2018, November 11). *ESA.* Tratto da ESA: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_L1C_Data_Quality_Report

European Commission. (2017, February 28). *Copernicus EU.* Retrieved September 08, 2017, from Copernicus Emergency Management Service: http://www.copernicus.eu/sites/default/files/documents/Copernicus_Factsheets/Copernicus_EmergencyMonitoring_Feb2017.pdf

European Space Agency (ESA). (2017, September 12). *ESA*. Tratto il giorno September 12, 2017 da ESA Sentinel Online: https://sentinel.esa.int/web/sentinel/missions/sentinel-2/mission-objectives

Fengming, H., Bing, X., Huabing, H., Qian, Y., & Peng, G. (2008). Modelling spatial-temporal change of Poyang Lake using multitemporal Landsat imagery. *International Journal of Remote Sensing, 29*(20), 5767-5784. doi:https://doi.org/10.1080/01431160802060912

Food and Agriculture Organization of the United Nations (FAO). (2015). *The impact of disasters on agriculture and food security.* Roma: FAO.

Food and Agriculture Organization of the United Nations (FAO). (2015). *The Impact of Natural Hazards and Disasters on Agriculture and Food Security and Nutrition.* Roma: FAO.

GISGeography. (2017, August 6). *100 Earth Shattering Remote Sensing Applications & Uses*. Retrieved September 11, 2017, from GISGeography: http://gisgeography.com/100-earth-remote-sensing-applications-uses/

Gomarasca, M. A. (2009). *Basics of Geomatics.* London: Springer. doi:10.1007/978-1-4020-9014-1

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, 211-221.

Gore, A. (1998). The Digital Earth - Understanding our planet in the 21st Century. *Australian Surveyor*, 0005-0326. doi:https://doi.org/10.1080/00050348.1998.10558728

Government of Canada. (2015, November 23). *Government of Canada*. Tratto il giorno September 11, 2017 da Natural Resources Canada: http://www.nrcan.gc.ca/node/9393

Haklay, M. (2010). How good is volunteered geographical information? A comparative study of OpenStreetMap and ordnance survey datasets. *Environment and Planning B: Planning and Design, 37*, 682–703.

Hansen, M. C., Potapov, P., M., M., M., H., Turubanova, S. A., & Tyukavina, A. (2013). High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 850-853. doi:10.1126/science.1244693

Huang, C., Chen, Y., Zhang, S., & Wu, J. (2018). Detecting, Extracting, and Monitoring Surface Water From Space Using Optical Sensors: A

Review. *Reviews of Geophysics*, 333-360. doi:https://doi.org/10.1029/2018RG000598

Humanitarian OpenStreetMap. (2013, January 27). *http://tasks.hotosm.org/*. Retrieved September 07, 2017, from Hot OSM: http://tasks.hotosm.org/about

INGV. (2017, Gennaio 01). *INGV*. Tratto il giorno Aprile 19, 2018 da INGV: http://cnt.rm.ingv.it/event/7073641

IOM. (2009). *Migration, Environment and Climate Change: Assesing the Evidence.* Geneva: Internation Organization for Migration.

Jain, S., Saraf, A., Goswami, G., & Ahmad, T. (2006). Flood inundation mapping using NOAA AVHRR data. *Water Resour Manage, 20*(6), 949–959. doi:https://doi.org/10.1007/s11269-006-9016-4

Lee, S. (2004). Application of logistic regression model and its validation for landslide susceptibility mapping using GIS and remote sensing data. *International Journal of Remote Sensing*, 1477-1491. doi:doi:10.1080/01431160412331331012

Li, W., Wang, S., & Li, J. (2014). Object based building extraction by QuickBird image for population estimation: A case study of the City of Waterloo. *IEEE Geoscience and Remote Sensing Symposium* (pp. 3176-3179). Quebec City: IGARSS.

Mbaye, L. M. (2017). Climate change, natural disasters, and migration. *IZA World of Labor* , 346 doi: 10.15185/izawol.346.

McFeeters, S. K. (1996). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing, 17*(7), 1425-1432. doi:10.1080/01431169608948714

Mumby, P. J., Skirving, W., Strong, A. E., T. Hardy, J., LeDrew, E. F., Hochberg, E. J., . . . David, L. T. (2004). Remote sensing of coral reefs and their physical environment. *Marine Pollution Bulletin*, 219-228. doi:https://doi.org/10.1016/j.marpolbul.2003.10.031

Munich Re. (2012). *2011 Natural catastrophes worldwide - Disaster Profiles.* Retrieved September 08, 2017, from Munich Re: http://www.preventionweb.net/files/24476_20120104munichre.pdf

NASA. (2017, August 04). *National Aerodinamics and Space Administration*. Retrieved September 11, 2017, from National Aerodinamics and Space Administration: https://www.nasa.gov/directorates/heo/scan/communications/outreach/funfacts/txt_passive_active.html

Neagoe, V., & Strugaru, G. (2008). A concurrent neural network model for pattern recognition in multispectral satellite imagery. *World Automation Congress* (pp. 1-6). Hawaii: IEEE.

OCHA/IDMC/NRC. (2009). *Monitoring disaster displacement in the context of climate change.* Geneva: United Nations Offi ce for Coordination of Humanitarian Affairs.

Open Data Commons. (2017, September 7). *https://opendatacommons.org/*. Retrieved September 07, 2017, from Open Data Commons: https://opendatacommons.org/licenses/odbl/

O'Reilly, T. (2005, September 09). *http://www.oreilly.com*. Retrieved September 07, 2017, from O'Reilly: http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Pekel, J.-F., Cottam, A., Gorelick, N., & S. Belward, A. (2016). High-resolution mapping of global surface water and its long-term changes. *Nature*(540), 418-422. doi:10.1038/nature20584

Plaza, A., & al., e. (2009, September). Recent advances in techniques for hyperspectral image processing. *Remote Sensing of Environment, 113*, S110-S122.

Save the Children. (2016). *Education Disrupted.* Singapore: Save the Children.

Shaker, I., Abd-Elrahman, A., K. Abdel-Gawad, A., & A Sherief, M. (2011). Building Extraction from High Resolution Space Images in High

Density Residential Areas in the Great Cairo Region. *Remote Sensing, 3*, 781-791.

Shen, L., & Li, C. (2010). Water body extraction from Landsat ETM+ imagery using adaboost algorithm. *2010 18th International Conference on Geoinformatics* (p. 1-4). Beijing: IEEE. doi:10.1109/GEOINFORMATICS.2010.5567762

Sinha, P., Verma, N. K., & Ayele, E. (2016). Urban Built-up Area Extraction and Change Detection of Adama. *International Journal of Advanced Remote Sensing and GIS*, 1886-1895.

Tiranni, G., & al, e. (2015). Scaling up to National/Regional Urban Extent Mapping Using Landsat Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 8*(7), 3710-3719.

Tomnod. (2015, June 30). *http://blog.tomnod.com*. Retrieved September 07, 2017, from Tomnod: http://blog.tomnod.com/improving-algorithms

Tomnod. (2016, March 23). *http://discourse.tomnod.com*. Retrieved September 07, 2017, from https://www.tomnod.com/: discourse.tomnod.com/t/adelaide-solar-panels/3232

United Nations. (2017). *The Sustainable Development Goals Report.* New York: United Nations publication issued by the Department of Economic and Social Affairs (DESA). Tratto da https://unstats.un.org/sdgs/files/report/2017/thesustainabledevelopmentgoalsreport2017.pdf

United States Geological Survey (USGS). (2016, November 30). *USGS science for a changing world*. Tratto il giorno September 08, 2017 da The USGS Land Cover Institute (LCI): https://landcover.usgs.gov/nalcms.php

United States Geological Survey (USGS). (2017, June 22). *Landsat Missions/Landsat project Description*. Retrieved September 12, 2017, from USGS: https://landsat.usgs.gov/landsat-project-description

UNOSAT. (2017, September 08). *UNITAR*. Tratto il giorno September 08, 2017 da UNOSAT: http://www.unitar.org/unosat/what-we-do

Voigt, S., Giulio-Tonolo, F., Lyons, J., Kucera, J., Jones, B., Schneiderhan, T., . . . al., e. (2016). Global trends in satellite-based emergency mapping. *Science, 353*(6296), 247–252.

W. Fritz, L. (1996). The Era of Commercial Earth Observation Satellites. *Photogrammetric Engineering and Remote Sensing, 62*(1), 39-45.

Wenbo, L., Zhiqiang, D., Feng, L., Dongbo, Z., Hailei, W., Yuanmiao, G., . . . Xiaoming, Z. (2013). A Comparison of Land Surface Water Mapping Using the Normalized Difference Water Index from TM, ETM+ and ALI. *Remote Sensing*, 5530-5549. doi:10.3390/rs5115530

Xinjian, S. L., & Yin, J. (2004). Extracting damaged building information from single remote sensing images of post-earthquake. *Geoscience and Remote Sensing Symposium IGARSS'04. Proceedings. 2004 IEEE International.* Anchorage: IEEE.

Xu, H. (2006). Modification of normalised difference water index (NDWI) to enhance open water features in remotely sensed imagery. *International Journal of Remote Sensing, 27*(14), 3025-3033. doi:https://doi.org/10.1080/01431160600589179

Yun, D., Yihang, Z., Feng, L., Qunming, W., Wenbo, L., & Xiaodong, L. (2016). Water Bodies' Mapping from Sentinel-2 Imagery with Modified Normalized Difference Water Index at 10-m Spatial Resolution Produced by Sharpening the SWIR Band. *Remote Sensing*, 354. doi:10.3390/rs8040354

Zahra, L., & Hamid, E. (2017). Automatic extraction of building features from high resolution satellite images using artificial neural networks. *ResearchGate*.

Zha, Y., Gao, J., & Ni, S. (2003). Use of normalized difference built-up index in automatically mapping. *International Journal of Remote Sensing, 24*(3), 583–594.

Zielstra, D., & Zipf, A. (2010). A Comparative Study of Proprietary Geodata and Volunteered Geographic Information for Germany. *13th AGILE International Conference on Geographic Information Science*.

Annexes

1_Beruwala — NDWI (-0.117)

		Predicted: Water	Predicted: Other	
Actual	Water	243818	3672	247490
	Other	11406	1763779	1775185
		255224	1767451	2022675

		Omission	Comission	PA	UA
Actual	Water	1,48%	4,47%	98,52%	95,53%
	Other	0,64%	0,21%	99,36%	99,79%
	Overall	99,25%			

1_Beruwala — MNDWI (0.242)

		Predicted: Water	Predicted: Other	
Actual	Water	243935	3555	247490
	Other	133604	1641580	1775184
		377539	1645135	2022674

		Omission	Comission	PA	UA
Actual	Water	1,44%	35,39%	98,56%	64,61%
	Other	7,53%	0,22%	92,47%	99,78%
	Overall	93,22%			

1_Beruwala — WRI (1.154)

		Predicted: Water	Predicted: Other	
Actual	Water	242929	4561	247490
	Other	4734	1770450	1775184
		247663	1775011	2022674

		Omission	Comission	PA	UA
Actual	Water	1,84%	1,91%	98,16%	98,09%
	Other	0,27%	0,26%	99,73%	99,74%
	Overall	99,54%			

1_Beruwala — thNDWI + thMNDWI (0.211)

		Predicted: Water	Predicted: Other	
Actual	Water	240849	6640	247489
	Other	5111	1770073	1775184
		245960	1776713	2022673

		Omission	Comission	PA	UA
Actual	Water	2,68%	2,08%	97,32%	97,92%
	Other	0,29%	0,37%	99,71%	99,63%
	Overall	99,42%			

1_Beruwala — thNDWI + thMNDWI + thWRI (0.211)

		Predicted: Water	Predicted: Other	
Actual	Water	240278	7211	247489
	Other	4195	1770989	1775184
		244473	1778200	2022673

		Omission	Comission	PA	UA
Actual	Water	2,91%	1,72%	97,09%	98,28%
	Other	0,24%	0,41%	99,76%	99,59%
	Overall	99,44%			

2_Viadana — NDWI (-0.332)

		Predicted: Water	Predicted: Other	
Actual	Water	38642	2276	40918
	Other	274563	466195	740758
		313205	468471	781676

		Omission	Comission	PA	UA
Actual	Water	5,56%	87,66%	94,44%	12,34%
	Other	37,07%	0,49%	62,93%	99,51%
	Overall	64,58%			

2_Viadana — MNDWI (0.226)

		Predicted: Water	Predicted: Other	
Actual	Water	31901	9017	40918
	Other	5535	735224	740759
		37436	744241	781677

		Omission	Comission	PA	UA
Actual	Water	22,04%	14,79%	77,96%	85,21%
	Other	0,75%	1,21%	99,25%	98,79%
	Overall	98,14%			

2_Viadana — WRI (1.147)

		Predicted: Water	Predicted: Other	
Actual	Water	30381	10537	40918
	Other	1568	739190	740758
		31949	749727	781676

		Omission	Comission	PA	UA
Actual	Water	25,75%	4,91%	74,25%	95,09%
	Other	0,21%	1,41%	99,79%	98,59%
	Overall	98,45%			

2_Viadana — thNDWI + thMNDWI (0.136)

		Predicted: Water	Predicted: Other	
Actual	Water	31383	9535	40918
	Other	2550	738208	740758
		33933	747743	781676

		Omission	Comission	PA	UA
Actual	Water	23,30%	7,51%	76,70%	92,49%
	Other	0,34%	1,28%	99,66%	98,72%
	Overall	98,45%			

2_Viadana — thNDWI + thMNDWI + thWRI (0.136)

		Predicted: Water	Predicted: Other	
Actual	Water	30119	10799	40918
	Other	1412	739346	740758
		31531	750145	781676

		Omission	Comission	PA	UA
Actual	Water	26,39%	4,48%	73,61%	95,52%
	Other	0,19%	1,44%	99,81%	98,56%
	Overall	98,44%			

3_Karunki — NDWI (-0.069)

		Predicted: Water	Predicted: Other	
Actual	Water	142677	6666	149343
	Other	7357	1145728	1153085
		150034	1152394	1302428

		Omission	Comission	PA	UA
Actual	Water	4,46%	4,90%	95,54%	95,10%
	Other	0,64%	0,58%	99,36%	99,42%
	Overall	98,92%			

3_Karunki — MNDWI (0.336)

		Predicted: Water	Predicted: Other	
Actual	Water	138827	10516	149343
	Other	6606	1146479	1153085
		145433	1156995	1302428

		Omission	Comission	PA	UA
Actual	Water	7,04%	4,54%	92,96%	95,46%
	Other	0,57%	0,91%	99,43%	99,09%
	Overall	98,69%			

3_Karunki — WRI (1.62)

		Predicted: Water	Predicted: Other	
Actual	Water	135001	14342	149343
	Other	3883	1149202	1153085
		138884	1163544	1302428

		Omission	Comission	PA	UA
Actual	Water	9,60%	2,80%	90,40%	97,20%
	Other	0,34%	1,23%	99,66%	98,77%
	Overall	98,60%			

3_Karunki — thNDWI + thMNDWI

		Predicted: Water	Predicted: Other	
Actual	Water	137941	11402	149343
	Other	5160	1147924	1153084
		143101	1159326	1302427

		Omission	Comission	PA	UA
Actual	Water	7,63%	3,61%	92,37%	96,39%
	Other	0,45%	0,98%	99,55%	99,02%
	Overall	98,73%			

3_Karunki — thNDWI + thMNDWI + thWRI

		Predicted: Water	Predicted: Other	
Actual	Water	135378	13964	149342
	Other	3989	1149095	1153084
		139367	1163059	1302426

		Omission	Comission	PA	UA
Actual	Water	9,35%	2,86%	90,65%	97,14%
	Other	0,35%	1,20%	99,65%	98,80%
	Overall	98,62%			

4_Puerto Antioquia — NDWI

		Predicted: Water	Predicted: Other	
Actual	Water	64747	3735	68482
	Other	66693	850237	916930
		131440	853972	985412

		Omission	Comission	PA	UA
Actual	Water	5,45%	50,74%	94,55%	49,26%
	Other	7,27%	0,44%	92,73%	99,56%
	Overall	92,85%			

4_Puerto Antioquia — MNDWI

		Predicted: Water	Predicted: Other	
Actual	Water	55347	13136	68483
	Other	17217	899713	916930
		72564	912849	985413

		Omission	Comission	PA	UA
Actual	Water	19,18%	23,73%	80,82%	76,27%
	Other	1,88%	1,44%	98,12%	98,56%
	Overall	96,92%			

4_Puerto Antioquia — WRI

		Predicted: Water	Predicted: Other	
Actual	Water	56374	9802	66176
	Other	7144	909786	916930
		63518	919588	983106

		Omission	Comission	PA	UA
Actual	Water	14,81%	11,25%	85,19%	88,75%
	Other	0,78%	1,07%	99,22%	98,93%
	Overall	98,28%			

4_Puerto Antioquia — thNDWI + thMNDWI

		Predicted: Water	Predicted: Other	
Actual	Water	53158	15324	68482
	Other	9552	907378	916930
		62710	922702	985412

		Omission	Comission	PA	UA
Actual	Water	22,38%	15,23%	77,62%	84,77%
	Other	1,04%	1,66%	98,96%	98,34%
	Overall	97,48%			

4_Puerto Antioquia — thNDWI + thMNDWI + thWRI

		Predicted: Water	Predicted: Other	
Actual	Water	51639	16842	68481
	Other	5755	911175	916930
		57394	928017	985411

		Omission	Comission	PA	UA
Actual	Water	24,59%	10,03%	75,41%	89,97%
	Other	0,63%	1,81%	99,37%	98,19%
	Overall	97,71%			

5_Magwitz — NDWI

		Predicted: Water	Predicted: Other	
Actual	Water	19416	1363	20779
	Other	99459	744283	843742
		118875	745646	864521

		Omission	Comission	PA	UA
Actual	Water	6,56%	83,67%	93,44%	16,33%
	Other	11,79%	0,18%	88,21%	99,82%
	Overall	88,34%			

5_Magwitz — MNDWI

		Predicted: Water	Predicted: Other	
Actual	Water	19915	863	20778
	Other	184353	659389	843742
		204268	660252	864520

		Omission	Comission	PA	UA
Actual	Water	4,15%	90,25%	95,85%	9,75%
	Other	21,85%	0,13%	78,15%	99,87%
	Overall	78,58%			

5_Magwitz — WRI

		Predicted: Water	Predicted: Other	
Actual	Water	17498	3281	20779
	Other	908	842833	843741
		18406	846114	864520

		Omission	Comission	PA	UA
Actual	Water	15,79%	4,93%	84,21%	95,07%
	Other	0,11%	0,39%	99,89%	99,61%
	Overall	99,52%			

5_Magwitz — thNDWI + thMNDWI

		Predicted: Water	Predicted: Other	
Actual	Water	19289	1490	20779
	Other	4010	839732	843742
		23299	841222	864521

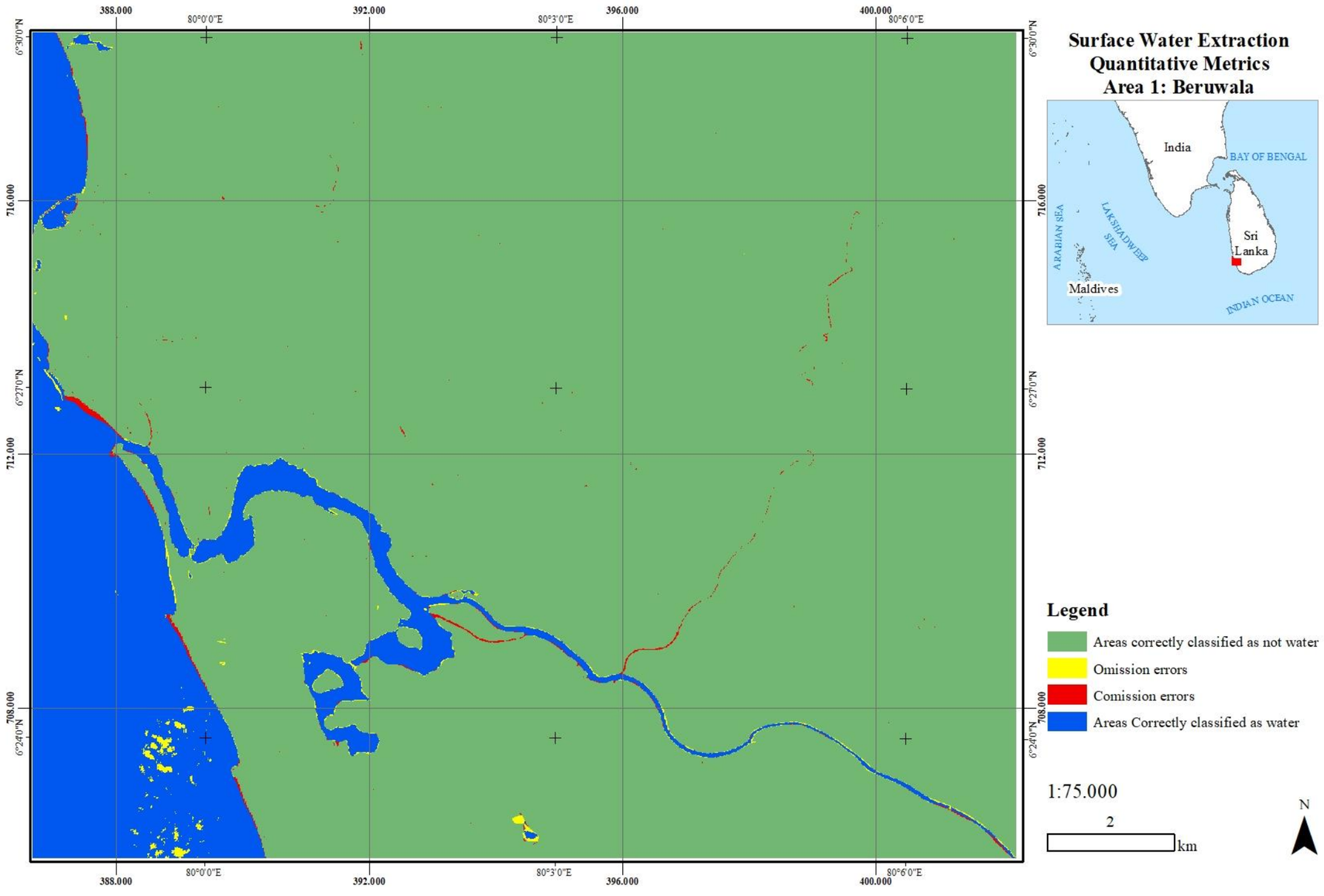
		Omission	Comission	PA	UA
Actual	Water	7,17%	17,21%	92,83%	82,79%
	Other	0,48%	0,18%	99,52%	99,82%
	Overall	99,36%			

5_Magwitz — thNDWI + thMNDWI + thWRI

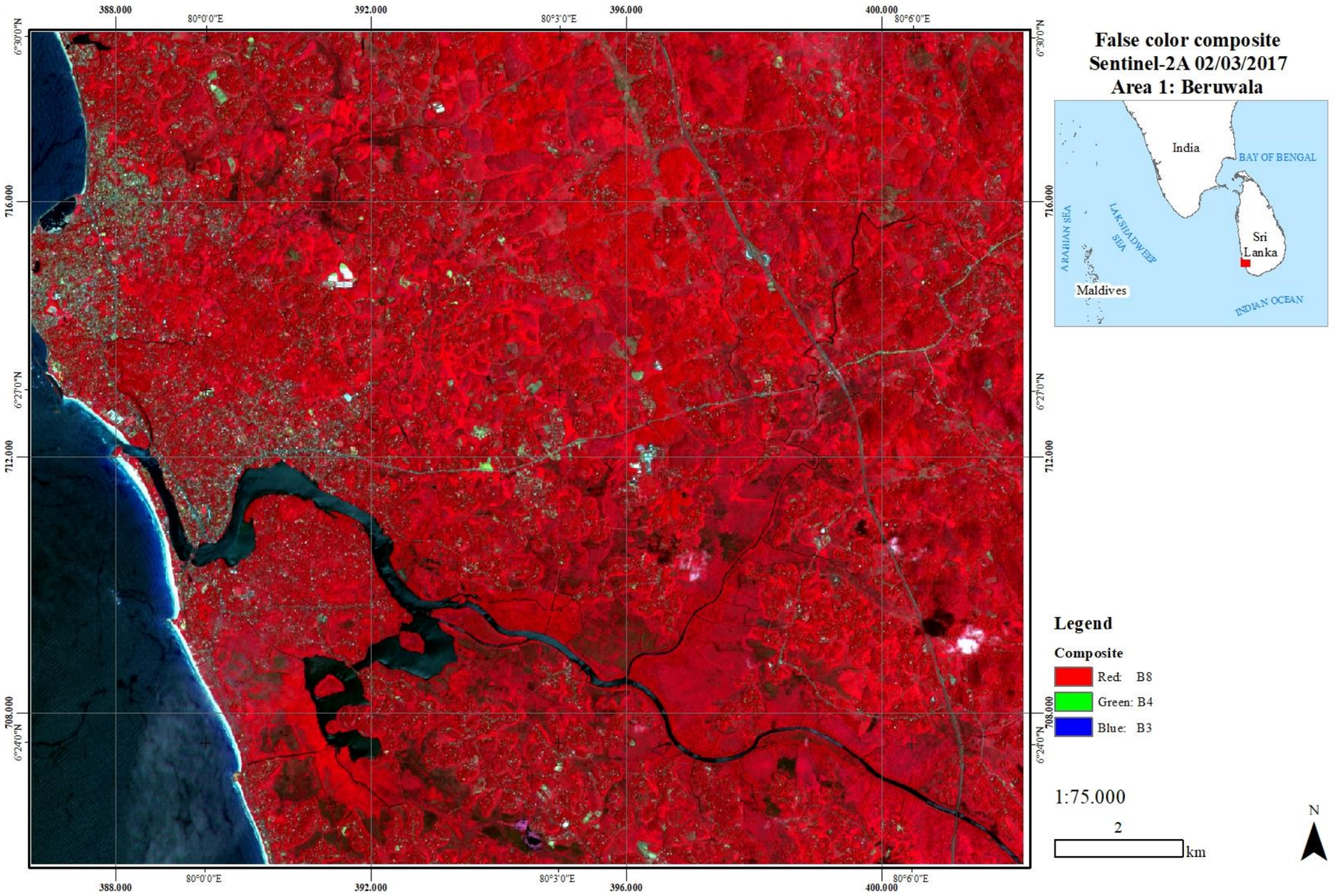
		Predicted: Water	Predicted: Other	
Actual	Water	17499	3280	20779
	Other	907	842834	843741
		18406	846114	864520

		Omission	Comission	PA	UA
Actual	Water	15,79%	4,93%	84,21%	95,07%
	Other	0,11%	0,39%	99,89%	99,61%
	Overall	99,52%			

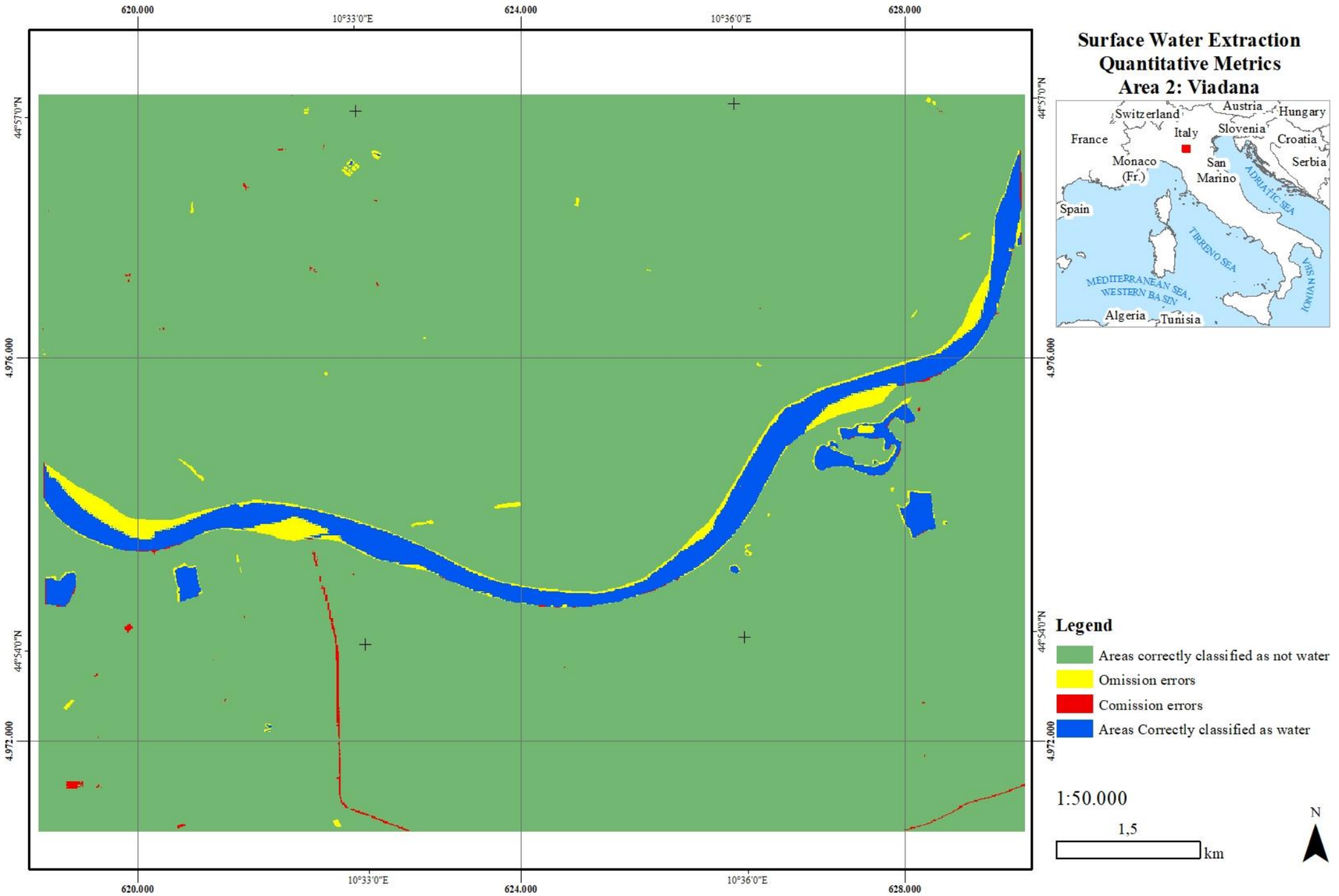
Annex 1 Detailed Error matrixes for each area of interest



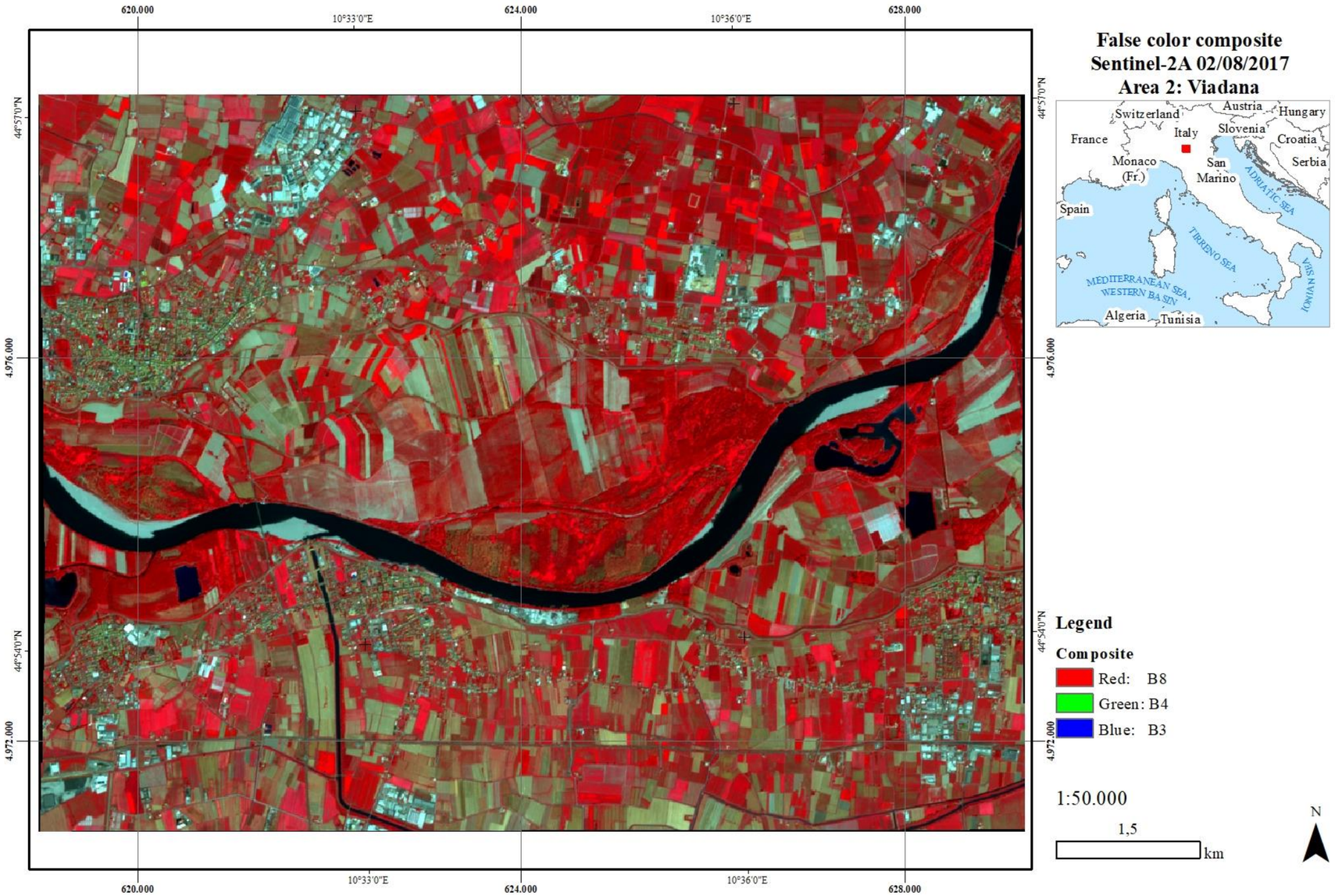
Annex 2 Area 1 - Surface Water Extraction Quantitative Metrics



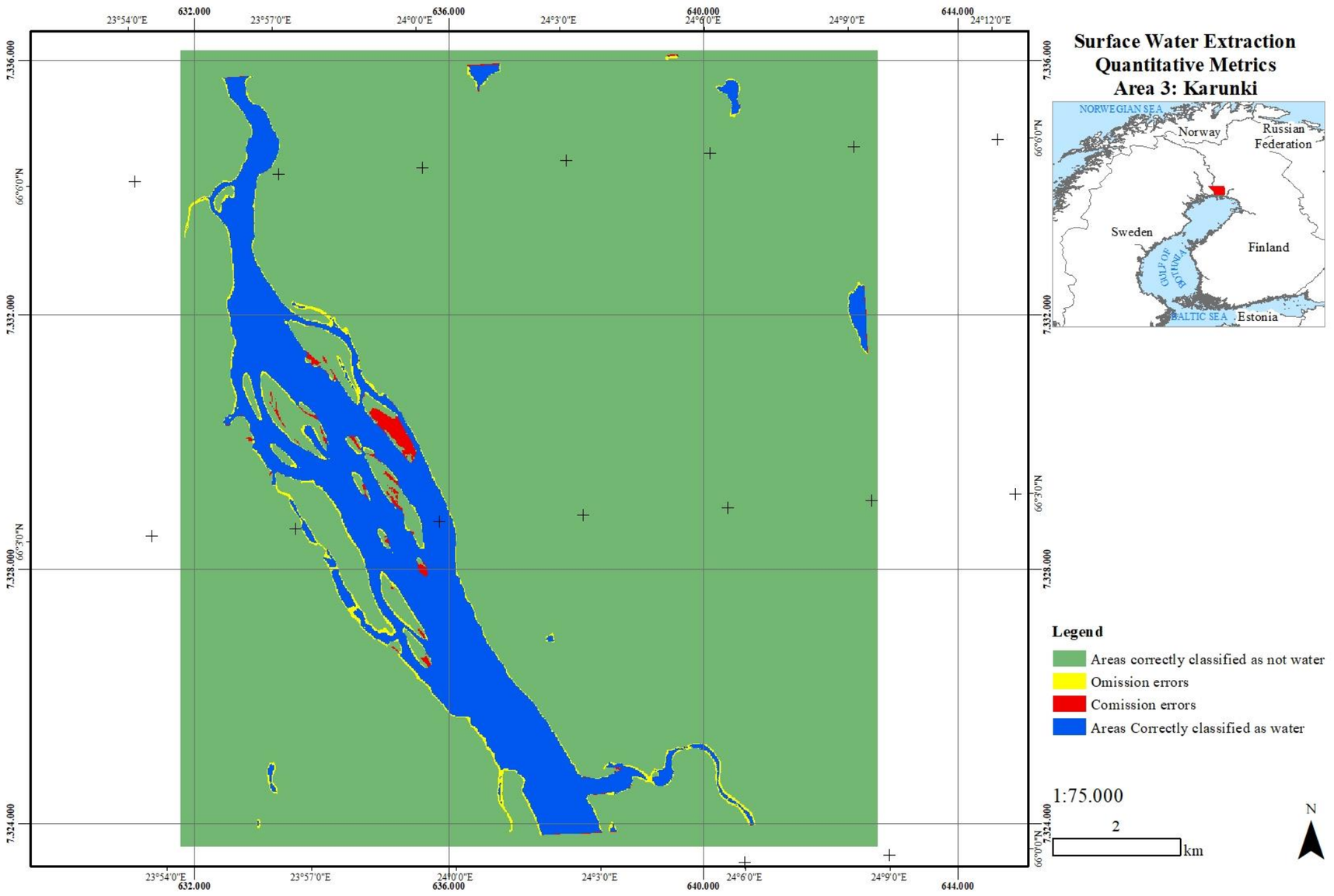
Annex 3 Area 1 - False colour composite (Bands: 8,4,3)



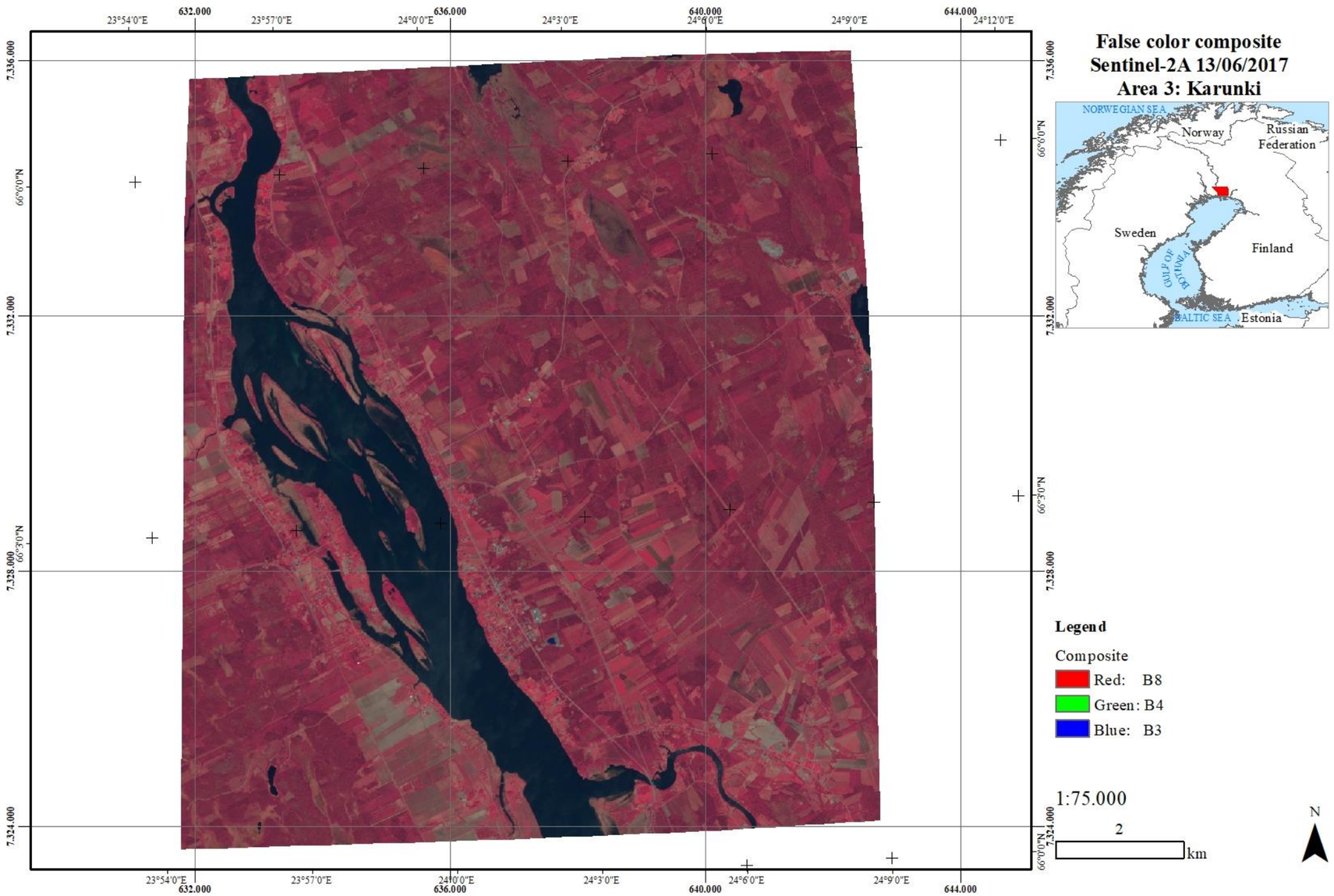
Annex 4 Area 2 - Surface Water Extraction Quantitative Metrics



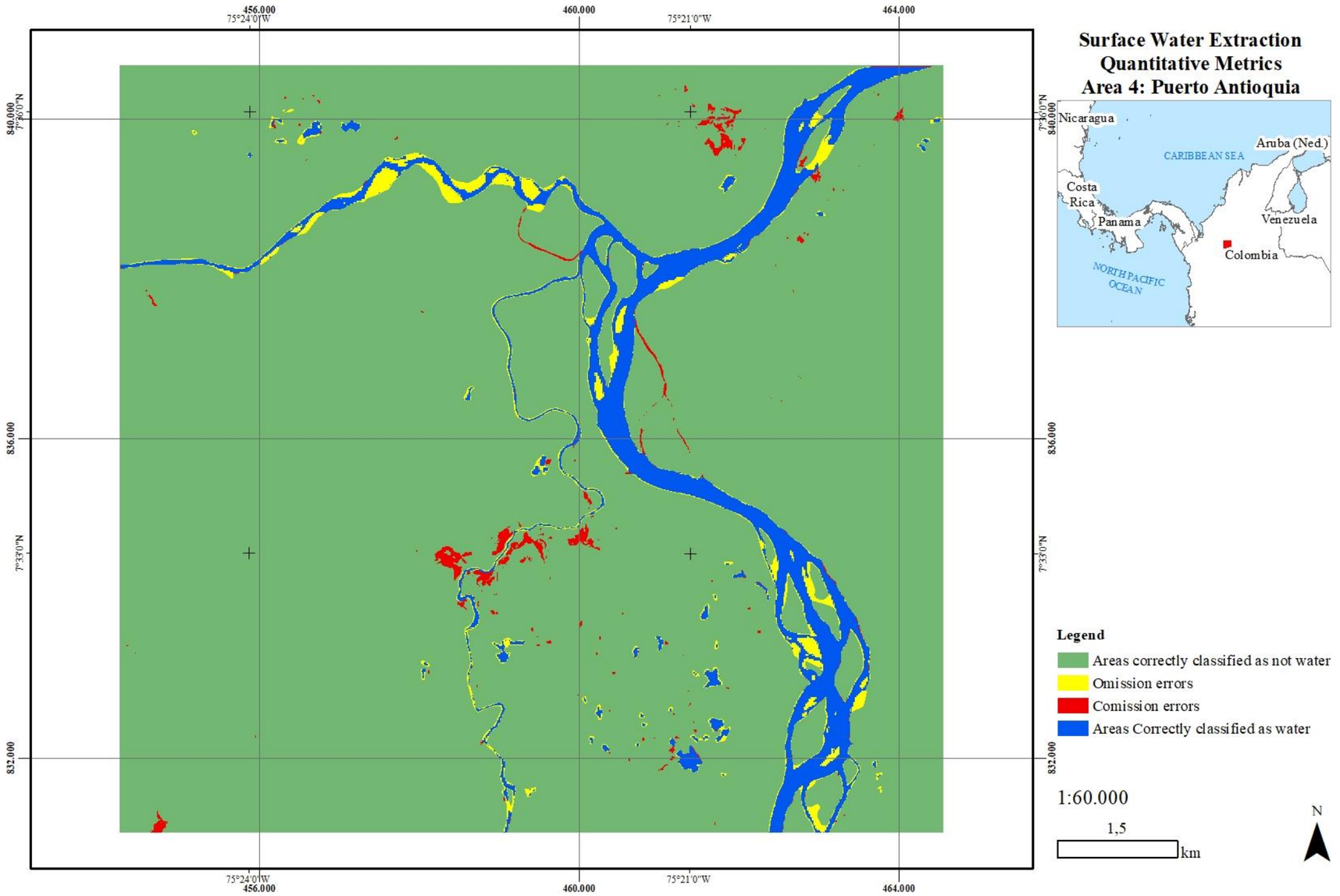
Annex 5 Area 2 - False colour composite (Bands: 8, 4, 3)



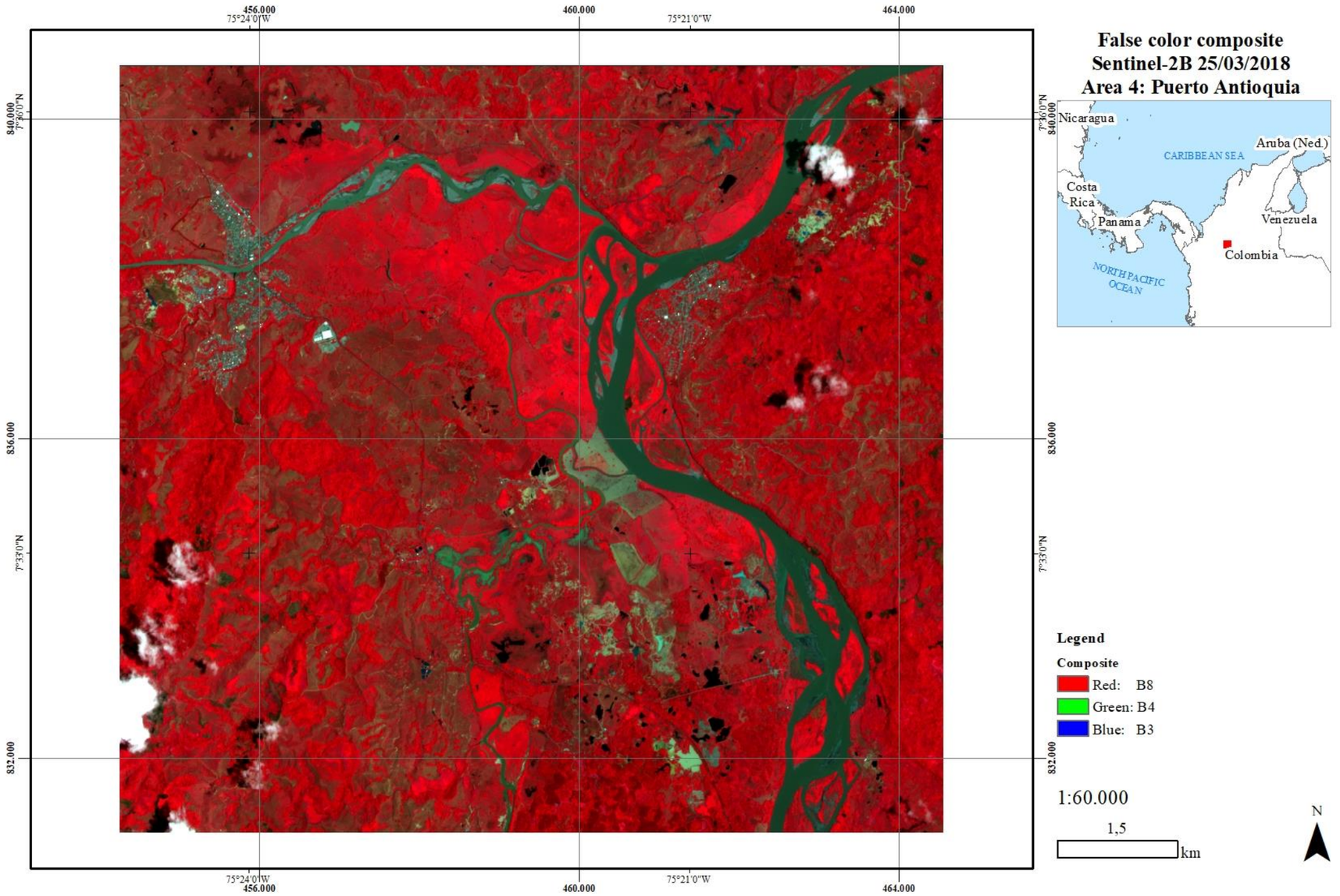
Annex 6 Area 3 - Surface Water Extraction Quantitative Metrics



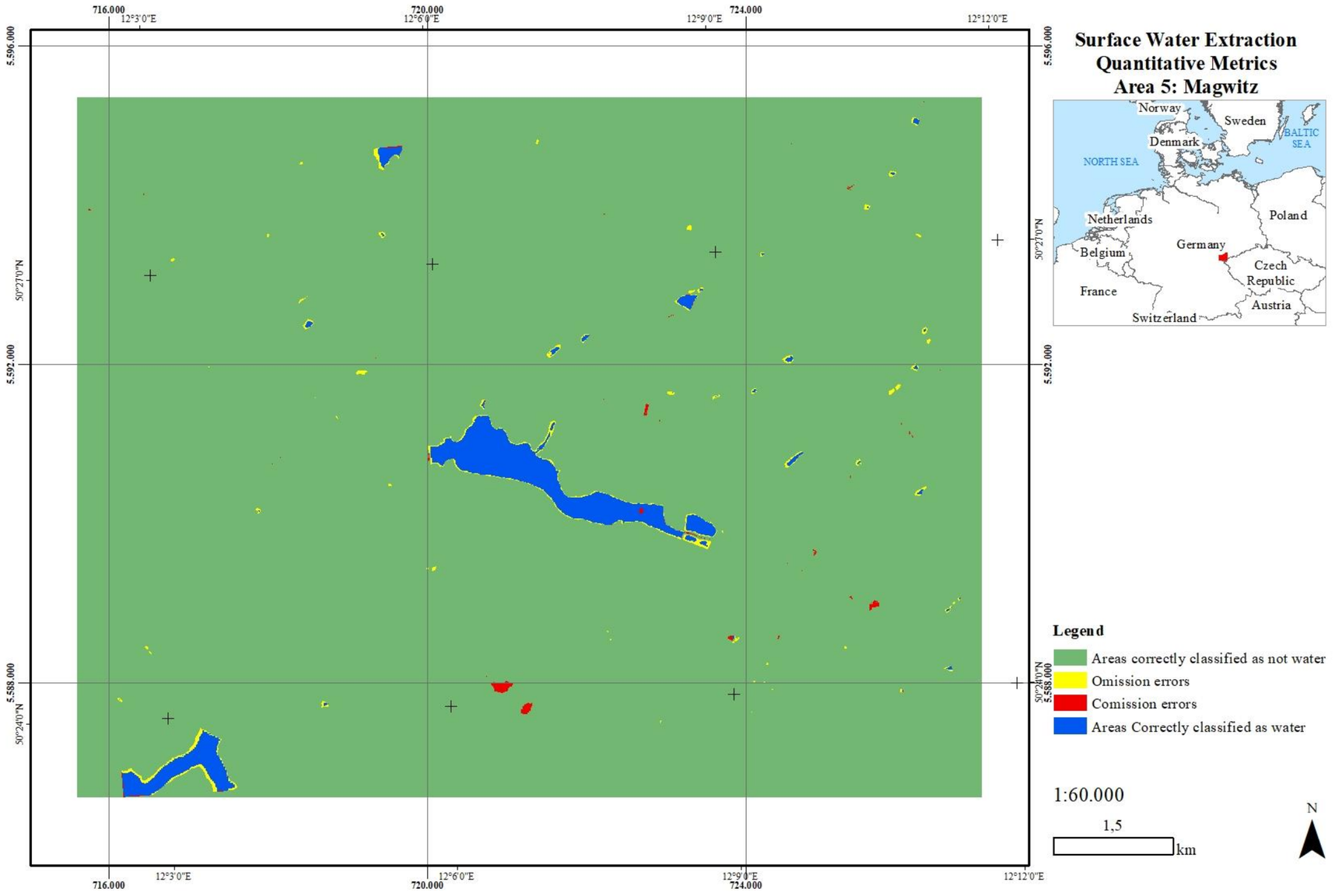
Annex 7 Area 3 - False colour composite (Bands: 8, 4, 3)



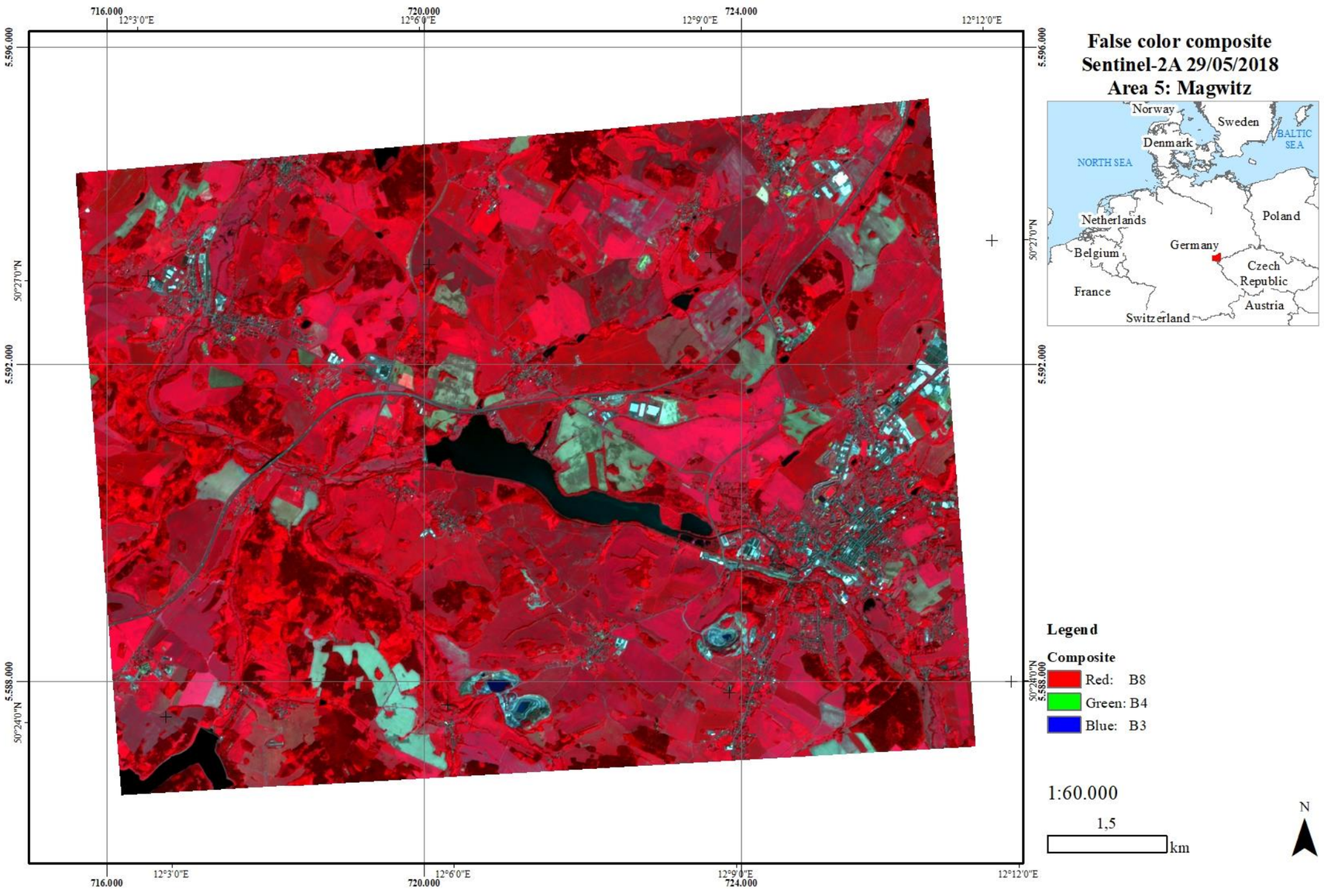
Annex 8 Area 4 - Surface Water Extraction Quantitative Metrics



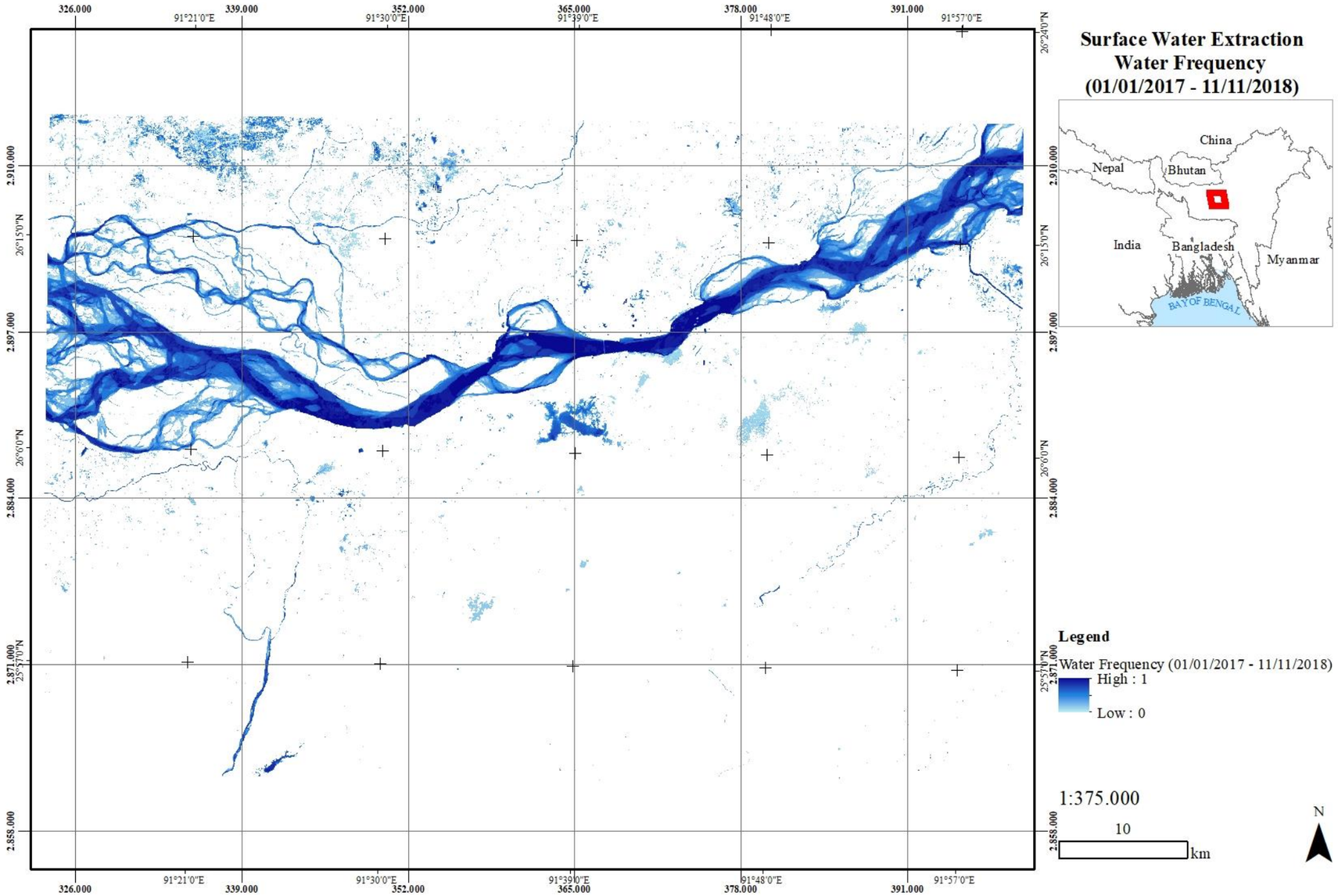
Annex 9 Area 4 - False colour composite (Bands: 8, 4, 3)



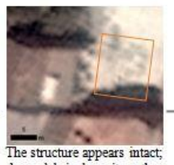
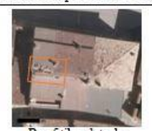
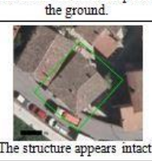
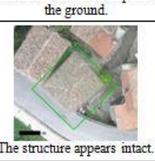
Annex 10 Area 5 - Surface Water Extraction Quantitative Metrics



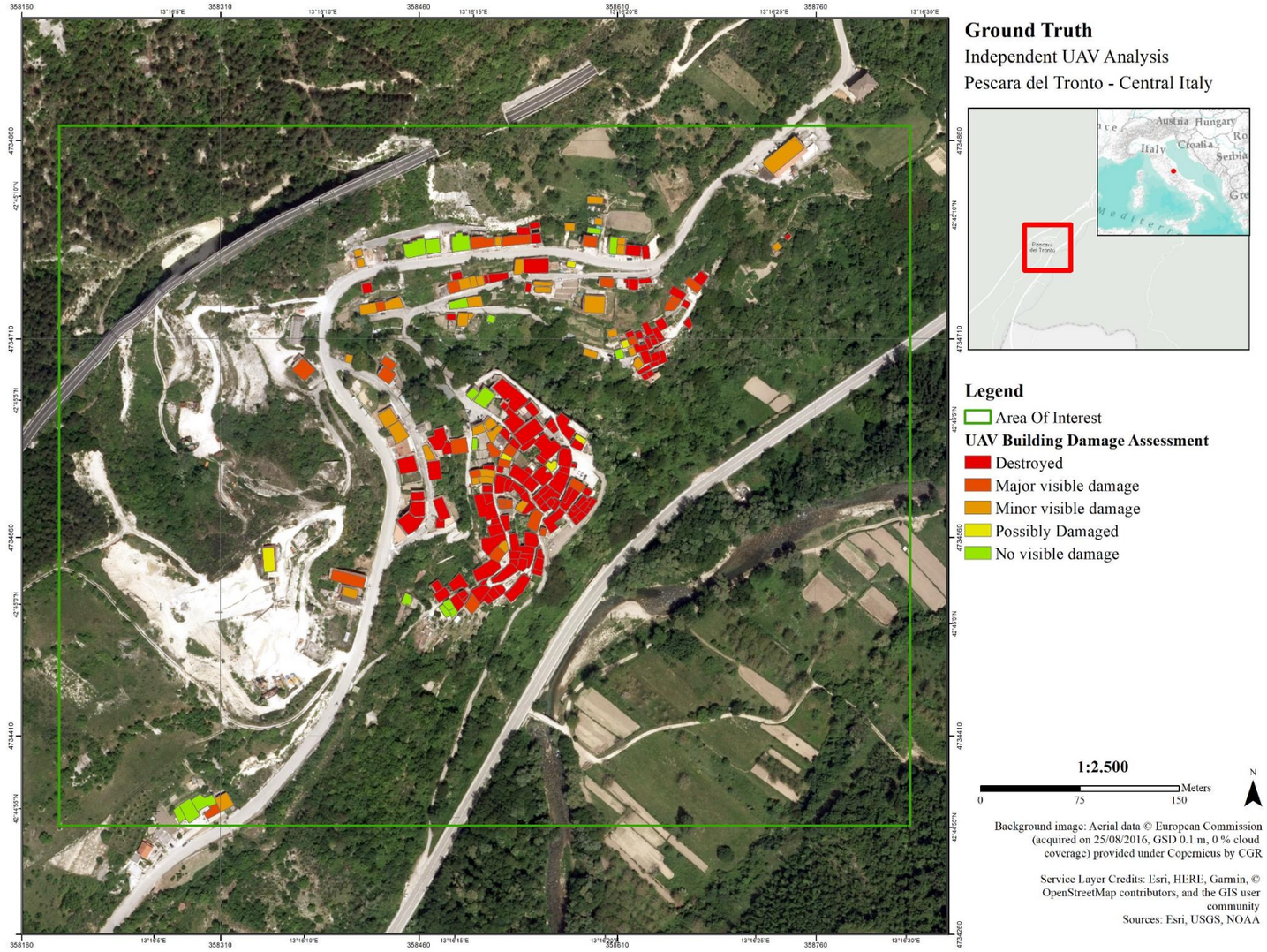
Annex 11 Area 5 - False colour composite (Bands: 8, 4, 3)



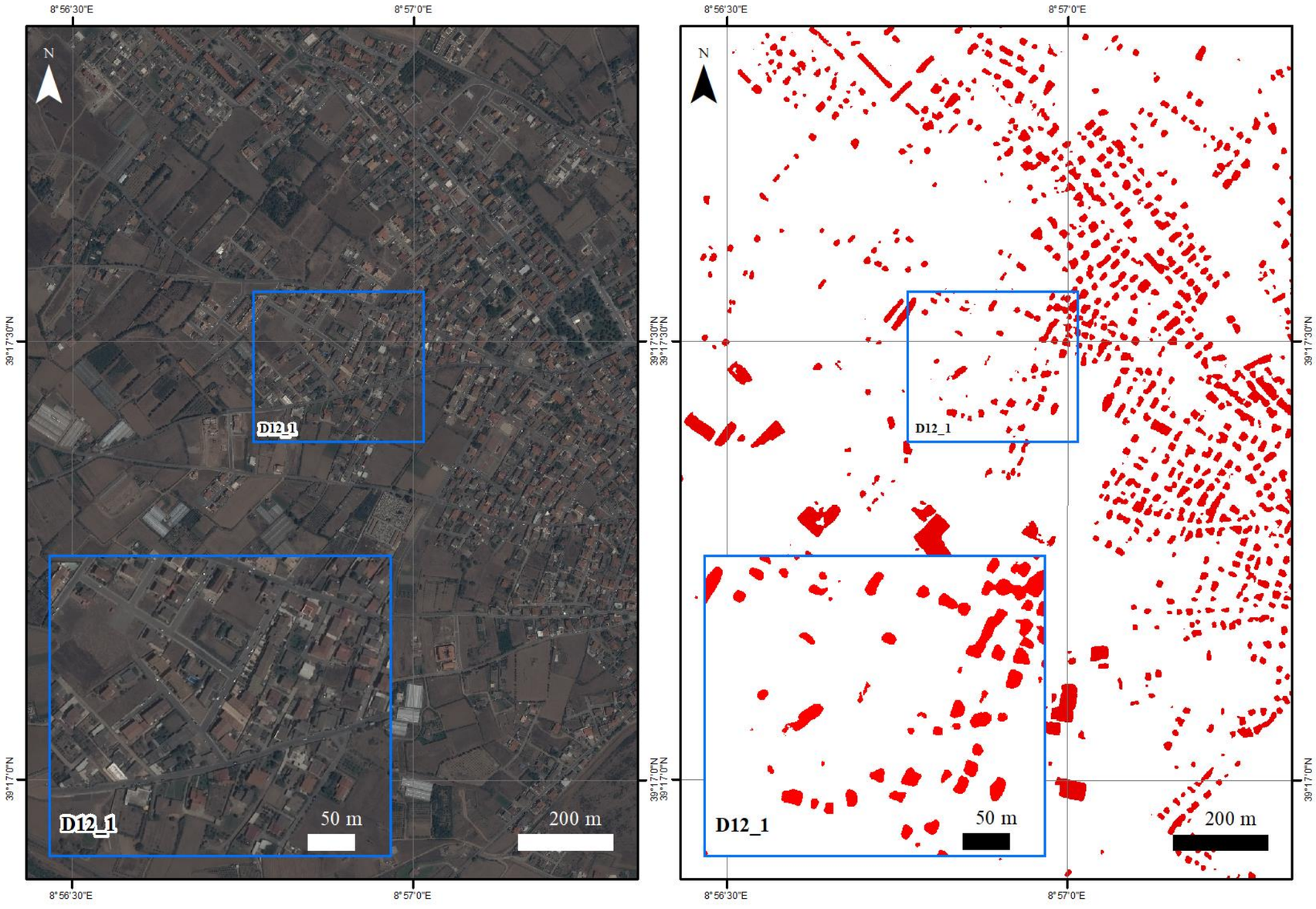
Annex 12 Water Frequency (01/01/2017 – 11/11/2018) near Guwahati

Damage Grade	Description of the related Damage	PRE Event Aerial (0.1m < GSD < 0.3m)	POST-Event imagery and brief Interpretation Guidelines: POST Event Satellite (0.3m < GSD < 0.5m)	POST Event Aerial (0.1m < GSD < 0.3m)	POST Event UAV (GSD < 0.1m)
Destroyed	Total collapse; collapse of part of the building (>50%); building structure not distinguishable (the walls have been destroyed or collapsed).		 Building structure not distinguishable.	 Total collapse of the building.	 Total collapse of the building.
Damaged "Major" and "Minor" shall be used only when Aerial/UAV imagery (GSD<0.3 m) is available. Only "Damaged" grade when using Satellite imagery.	**Major**: Partial collapse of the roof; serious failure of walls (Tip: black spots on the rooftop suggest collapse of part of the roof).		Only "Damaged" grade when using Satellite imagery.  Black spots on the rooftop suggest collapse of part of the roof.	 Partial collapse of the roof.	 Partial collapse of the roof
	Minor: The roof remains largely intact, but presents partial damage (Tip: white spots on the rooftop suggest tiles' lack or displacement and collapse of chimneys). Presence of damage proxies like large debris/rubble or sand deposit around the building.	 	 The structure appears intact; large debris deposit on the ground.	 Roof tiles detach.  The structure appears intact; large debris deposit	 Collapse of chimneys, roof tiles detach.  The structure appears intact; large debris deposit
Possibly Damaged	Uncertain interpretation due to image quality (e.g. shadow or degraded resolution due to high off-nadir angle). Presence of possible damage proxies like small traces of debris/rubble or sand deposit around building. Building surrounded by damaged/destroyed buildings.		 Small traces of debris deposit on the ground.	 Small traces of debris deposit on the ground.	 Small traces of debris deposit on the ground.
No Visible Damage	The structure appears to have complete structural integrity; the walls remain standing; the roof is virtually undamaged. **The building may anyway suffered damages that can not be assessed from vertical satellite imagery (from no to slight structural damage).**		 The structure appears intact.	 The structure appears intact.	 The structure appears intact.

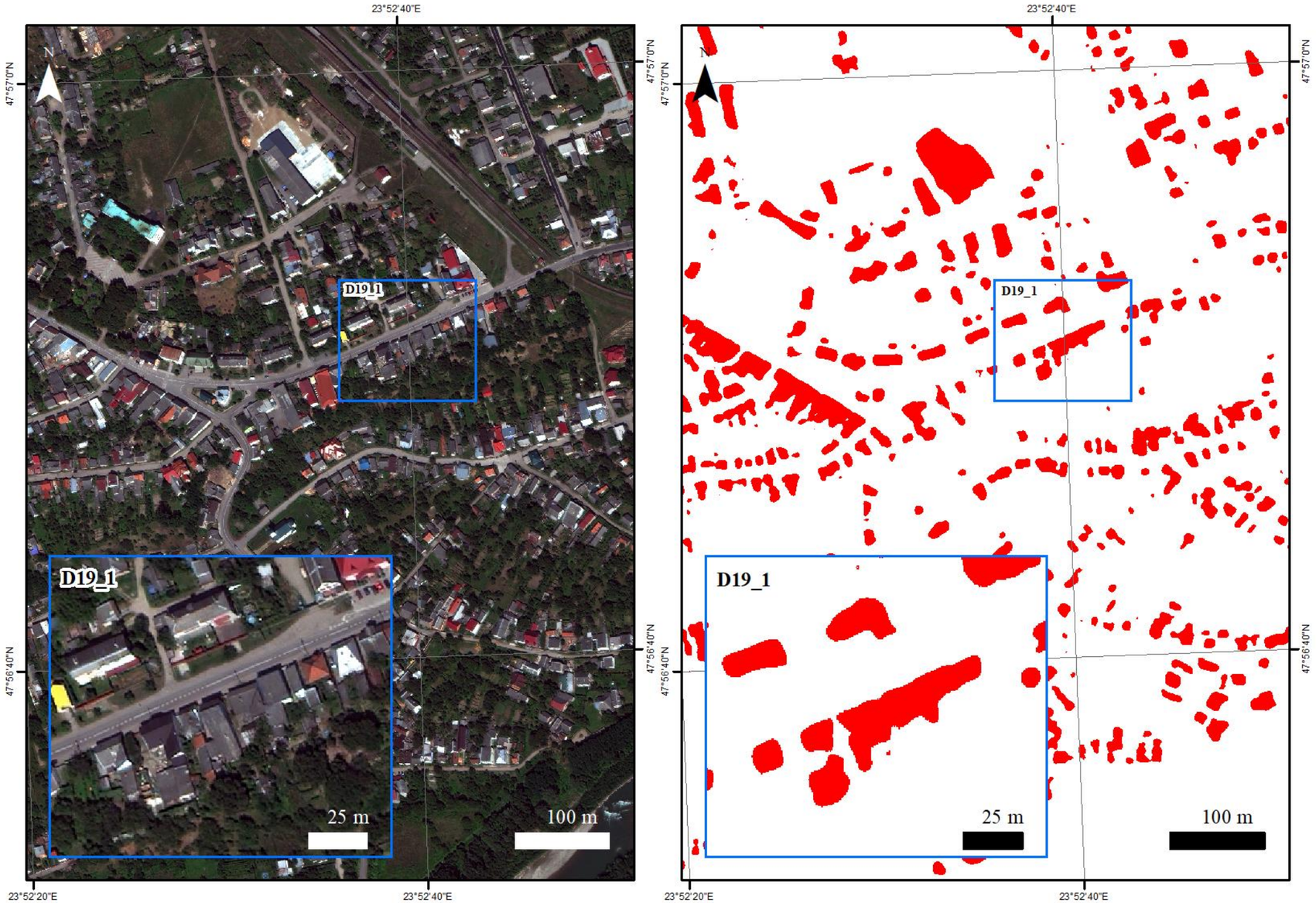
Annex 13 Proposed Building Damage Scale



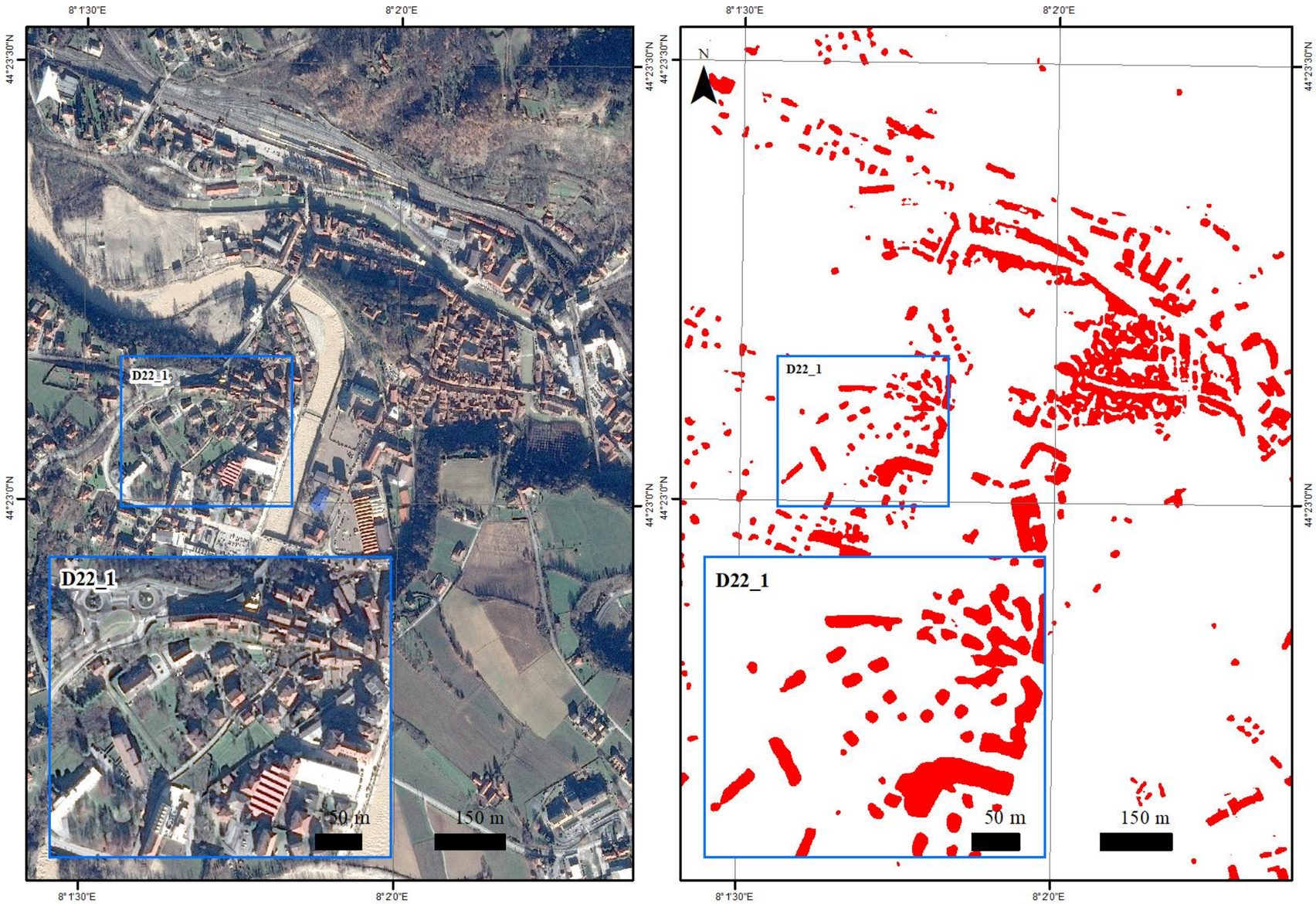
Annex 14 Pescara del Tronto Damage Assessment using the new Proposed Damage Scale



Annex 15 Detail of D12 building footprint extraction

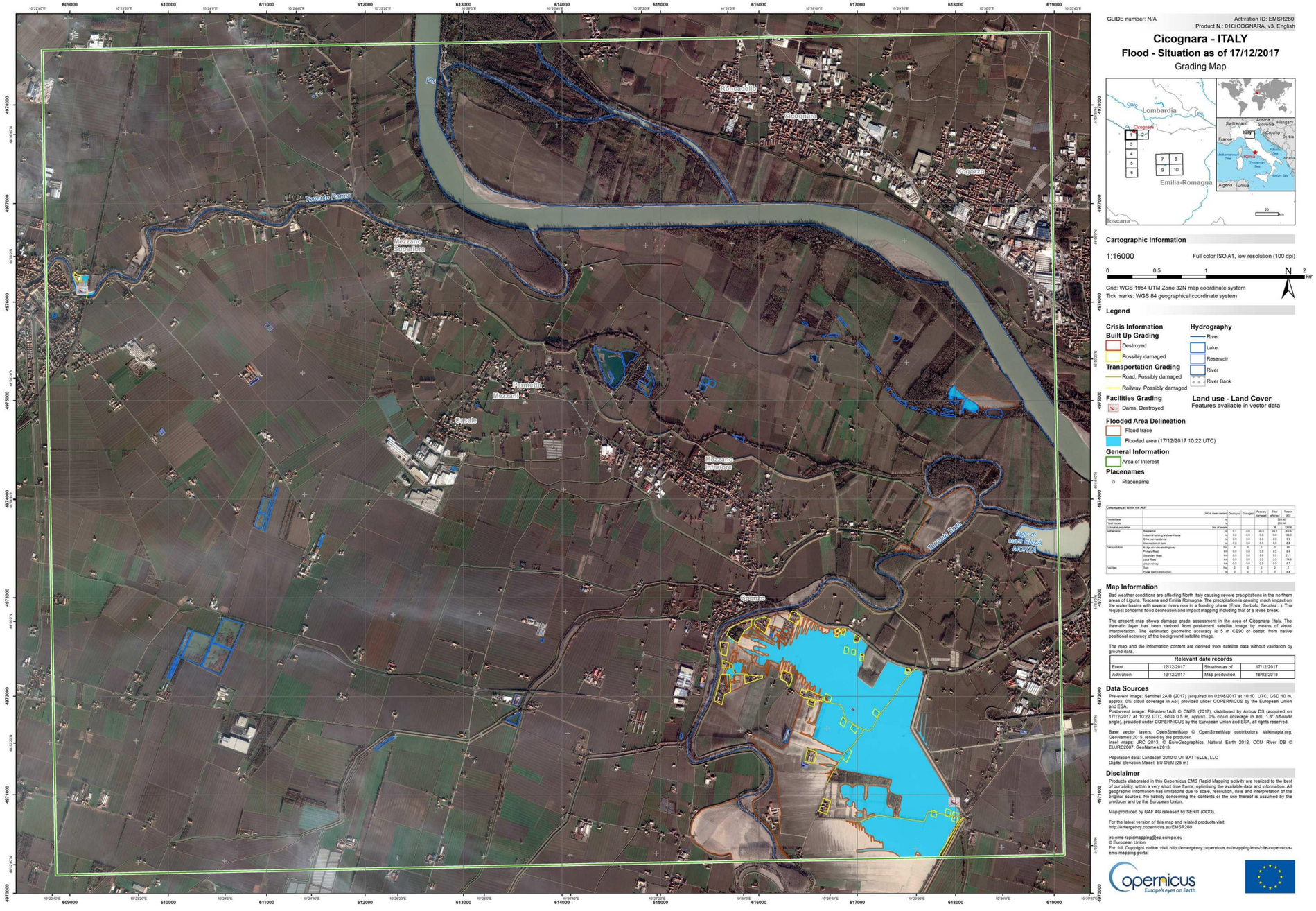


Annex 16 Detail of D19 building footprint extraction



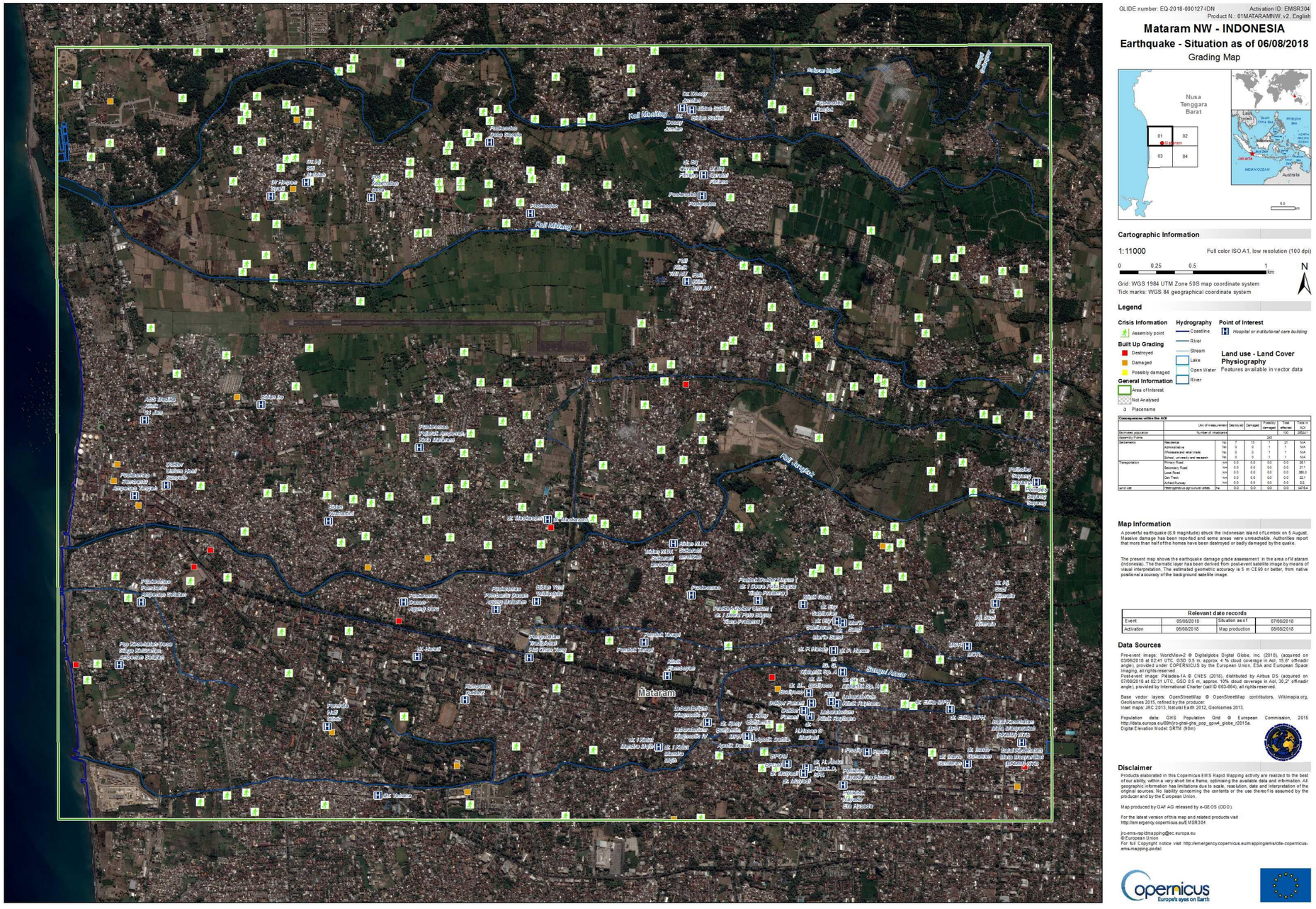
Annex 17 Detail of D22 building footprint extraction

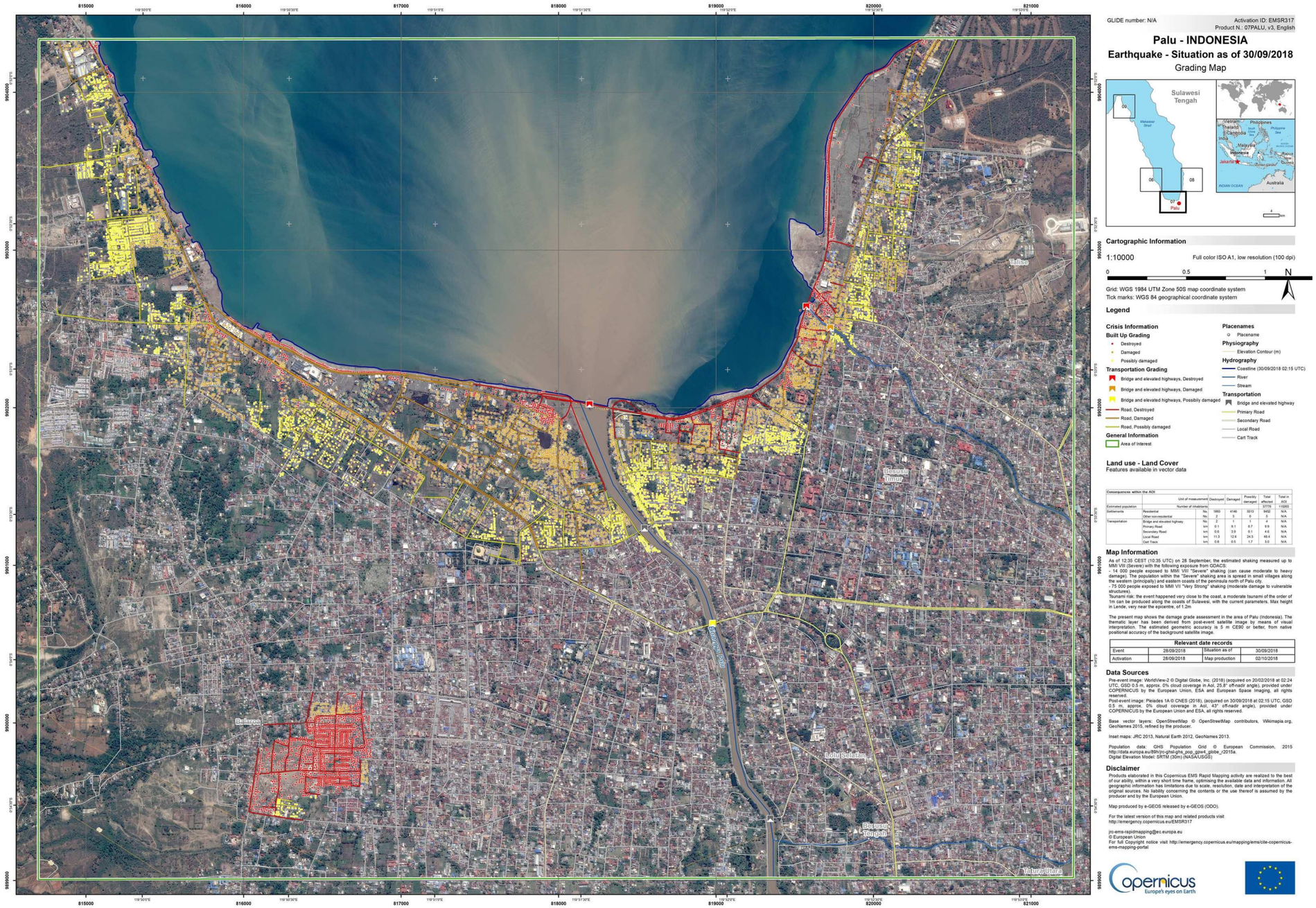
Annex 18 [EMSR257] Mandra: Grading Map (November 2017)



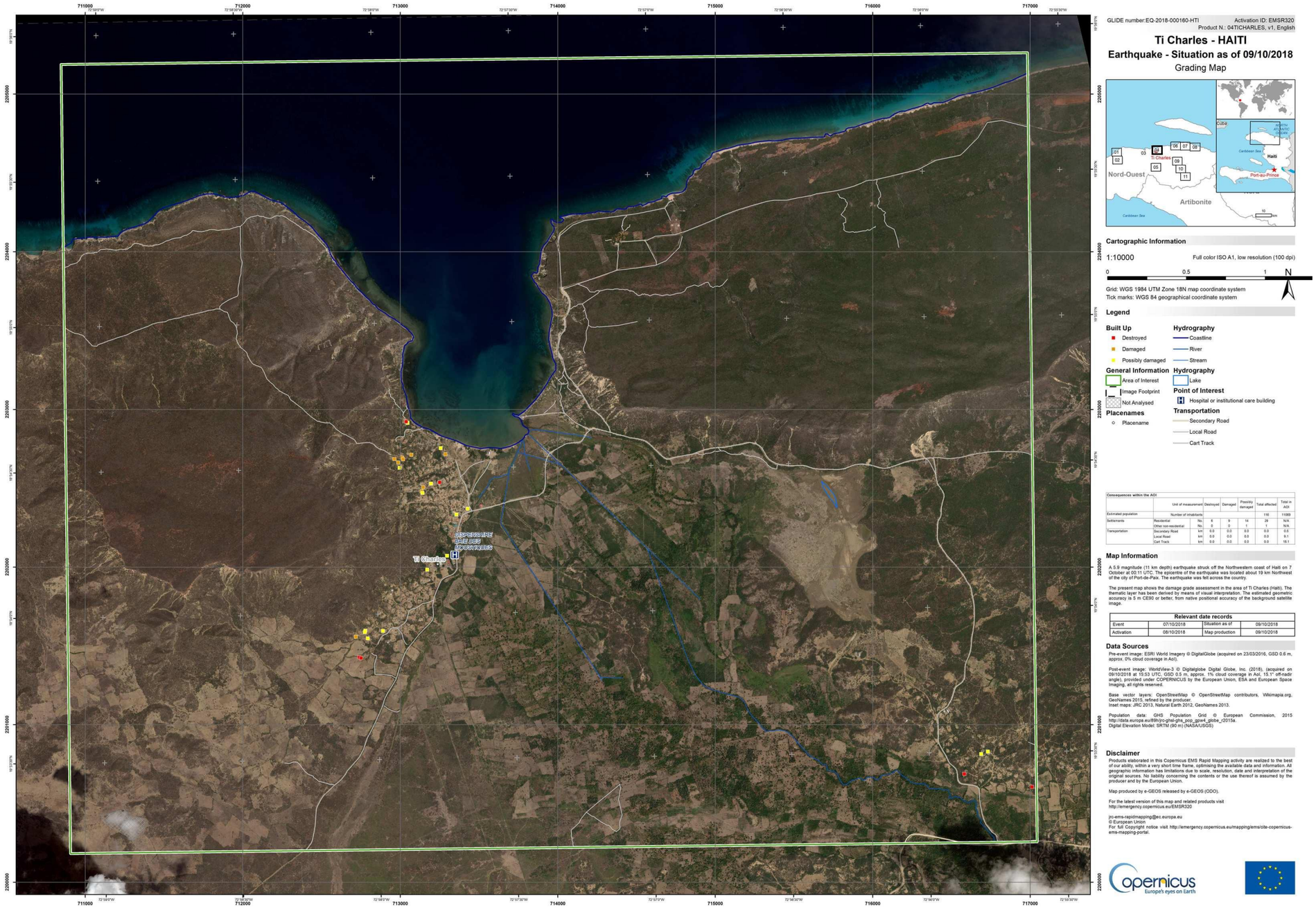
Annex 19 [EMSR260] Cicognara: Grading Map (December 2017)

Annex 20 [EMSR269] Nuku'alofa: Grading Map (February 2018)


Mataram
Rail Meulang
Rail Mataram
Rail Langko
Rail Kebon Roek
Rail Ampenan
Rail Kebon Roek
Rail Ampenan
Rail Mataram
Rail Langko



Annex 22 [EMSR317] Palu: Grading Map (September 2018)



Annex 23 [EMSR320] Ti Charles: Grading Map (October 2018)

1_Code used for the validation process

```
   /*The defined AOI choosed from Copernicus RM activations*/
   var geometryBeruwala = /* color: #3ad608 */ee.Geometry.Polygon(
           [[[79.97547987248322, 6.382708327515678],
             [80.11591558042426, 6.382976877368193],
             [80.11571156287789, 6.500765393907194],
             [79.97525446617942, 6.5004886695111123]]]),
       imageVisParam = {"opacity":1,"bands":["B8","B3","B2"],"min":-
239.8078626099284,"max":2572.733573015612,"gamma":1},
       imageVisParam2 = {"opacity":1,"bands":["B12","B8","B2"],"min":-
239.8078626099284,"max":2572.733573015612,"gamma":1},
       imageVisParam3 =
{"opacity":1,"bands":["B8","B4","B3"],"min":189.73812687529698,"max":256
0.190492653323,"gamma":1},
       samplingPoints = /* color: #35e6ef */ee.Geometry.MultiPoint(
           [[10.597185376538505, 44.90616148047311],
            [24.00982548776858, 66.04810337592632],
            [-75.34496759418386, 7.5722516681259755],
            [12.1029871525933, 50.42457945657212],
            [80.02208773490611, 6.421394523250352]]),
       AOIStatRegion = /* color: #d63000 */ee.Geometry.Point([-
75.34534734147672, 7.572445287392769]),
       imageVisParam4 =
{"opacity":1,"bands":["CL"],"max":4,"palette":["19d60d","ff0000","ff0000
","19d60d"]},
       geometryViadana = /* color: #a81dd6 */ee.Geometry.Polygon(
           [[[10.508331701872521, 44.95210709867036],
             [10.508090209456782, 44.91837129046192],
             [10.507961463424067, 44.91416226904162],
             [10.508050167392867, 44.906849135975875],
             [10.506738567959019, 44.88301575081166],
```

```
            [10.636508019020994, 44.88142683889363],
            [10.637667847155853, 44.935642851865744],
            [10.637916226865286, 44.950100079392311]]]),
    geometryKarunki = /* color: #98ff00 */ee.Geometry.Polygon(
        [[[23.91670533807701, 66.09477923148471],
          [23.90528332429608, 66.00604897158112],
          [24.01649540372273, 66.00530127836480],
          [24.08138369843664, 66.00510059523928],
          [24.104472217092734, 66.00517040232992],
          [24.14670103743606, 66.00541470462252],
          [24.15144089040257, 66.0676457493629],
          [24.15135505963508, 66.08066434139823],
          [24.149845296592275, 66.11360290655142],
          [24.093085757173185, 66.11414946065366],
          [23.920572387637208, 66.11398876082468]]]),
    geometryPuerto = /* color: #0b4a8b */ee.Geometry.Polygon(
        [[[-75.32138649332006, 7.605155736546275],
          [-75.41459916807605, 7.605071091111962],
          [-75.41450195835398, 7.5183886723325974],
          [-75.32127851314704, 7.5184552955469881]]]),
    geometryMag = /* color: #ffc82d */ee.Geometry.Polygon(
        [[[12.041577671393952, 50.39127621323662],
          [12.193264022312178, 50.392732957226265],
          [12.189831675365667, 50.4663340608964],
          [12.13081971670158, 50.464343331346186],
          [12.03795091176994, 50.46188463676546]]]),
    gTru = {"opacity":1,"bands":["b1"],"min":-
0.20481570297220456,"max":0.44312447939757293,"gamma":1},
```

```
      validationVisParam =
{"opacity":1,"bands":["MNDWI"],"max":4,"palette":["19ff15","ffa434","ff0
000","0726ff"]};


      var regionOfInt = AOIStatRegion.buffer(2000);


      //Images for validation

      var Beruwala_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth').clip(geometryBeruwala);


      var S2_Viadana =
ee.Image('COPERNICUS/S2/20170802T101031_20170802T101051_T32TPQ').clip(ge
ometryViadana);

      var Viadana_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth_Viadana').clip(geometryViada
na);


      var S2_Karunki =
ee.Image('COPERNICUS/S2/20170613T101031_20170613T101025_T34WFU').clip(ge
ometryKarunki);

      var Karunki_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth_Karunki').clip(geometryKarun
ki);


      var S2_Puerto =
ee.Image('COPERNICUS/S2/20180325T152639_20180325T152951_T18NVP').clip(ge
ometryPuerto);

      var Puerto_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth_Puerto').clip(geometryPuerto
);


      var S2_Puerto =
ee.Image('COPERNICUS/S2/20180325T152639_20180325T152951_T18NVP').clip(ge
ometryPuerto);

      var Puerto_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth_Puerto').clip(geometryPuerto
);
```

```
      var S2_Mag =
ee.Image('COPERNICUS/S2/20180529T101031_20180529T101225_T32UQA').clip(ge
ometryMag);

      //var S2_Mag =
ee.Image('COPERNICUS/S2/20170213T101121_20170213T101553_T32UQA').clip(ge
ometryMag);

      var Mag_GroudTruth =
ee.Image('users/consandu/EE_01_ground_truth_Mag').clip(geometryMag);



      //SWE function used for validation in this case we also plot the
graphs to better comprehend the distribution of the DN

      var SWE = function(S2_image, region){

        var date = ee.Date(S2_image.get('system:time_start'));

        var months = date.difference(ee.Date('1970-01-01'), 'year');

        var WRI = S2_image.select('B3').add(S2_image.select('B4'))


.divide(S2_image.select('B8').add(S2_image.select('B12'))).rename('WRI')
;//the final index

        var NDWI =  S2_image.normalizedDifference(['B3',
'B8']).rename('NDWI'); //McFeeters

        var MNDWI =  S2_image.normalizedDifference(['B3',
'B12']).rename('MNDWI'); //Xu but with the Swir not MIR (also present on
SNAP)

        var sintIm = ee.Image([WRI, NDWI, MNDWI]);

        var histogram_WRI =
ui.Chart.image.histogram(sintIm.select('WRI'), region,
30).setSeriesNames(['WRI']).setOptions({

        title: 'WRI',

        fontSize: 20,

        hAxis: {title: 'DN distribution'},

        vAxis: {title: 'Count of DN'},

        series: {

          0: {color: 'blue'}
```

```
          //1: {color: 'green'},
          //2: {color: 'red'}}
        }});
        var histogram_NDWI =
ui.Chart.image.histogram(sintIm.select('NDWI'), region,
30).setSeriesNames(['NDWI']).setOptions({
        title: 'NDWI',
        fontSize: 20,
        hAxis: {title: 'DN distribution'},
        vAxis: {title: 'Count of DN'},
        series: {
          0: {color: 'green'}
          //1: {color: 'green'},
          //2: {color: 'red'}}
        }});
        var histogram_MNDWI =
ui.Chart.image.histogram(sintIm.select('MNDWI'), region,
30).setSeriesNames(['MNDWI']).setOptions({
        title: 'MNDWI',
        fontSize: 20,
        hAxis: {title: 'DN distribution'},
        vAxis: {title: 'Count of DN'},
        series: {
          0: {color: 'red'}
          //1: {color: 'green'},
          //2: {color: 'red'}}
        }});
        var histogram_ALL = ui.Chart.image.histogram(sintIm, region,
30).setSeriesNames(['WRI', 'NDWI', 'MNDWI']).setOptions({
        title: 'Superimposition of the analyzed indexes',
        fontSize: 20,
```

```
        hAxis: {title: 'DN distribution'},
        vAxis: {title: 'Count of DN'},
        series: {
          0: {color: 'blue'},
          1: {color: 'green'},
          2: {color: 'red'},
        }});
        var histogram_NDWI_MNDWI =
ui.Chart.image.histogram(sintIm.select(['NDWI', 'MNDWI']), region,
30).setSeriesNames(['NDWI', 'MNDWI']).setOptions({
        title: 'Superimposition of the NDWI and MNDWI indexes',
        fontSize: 20,
        hAxis: {title: 'DN distribution'},
        vAxis: {title: 'Count of DN'},
        series: {
          0: {color: 'red'},
          1: {color: 'green'}
          //2: {color: 'blue'},
        }});
        print(histogram_WRI);
        print(histogram_NDWI);
        print(histogram_MNDWI);
        print(histogram_ALL);
        print(histogram_NDWI_MNDWI);
        //var fIndex = mWRI.add(NDWI).add(mNDWI2).rename('CL');
        //var minusMwri =
NDWI.add(MNDWI).subtract(mWRI).rename('CL');//.addBands(ee.Image(months)
.rename('time'));
        //var addMwri =
NDWI.add(MNDWI).add(mWRI).rename('CL');//.addBands(ee.Image(months).rena
me('time'));
```

```
        var addWri =
NDWI.add(MNDWI).add(WRI).rename('CL');//.addBands(ee.Image(months).renam
e('time'));

        var fIndex =
NDWI.add(MNDWI).rename('CL');//.addBands(ee.Image(months).rename('time')
);

        var histogramfIndex =
ui.Chart.image.histogram(fIndex.select('CL'), region,
30).setSeriesNames(['fIndex']).setOptions({

        title: 'Sum of WRI, NDWI, MNDWI',

        fontSize: 20,

        hAxis: {title: 'DN distribution'},

        vAxis: {title: 'Count of DN'},

        series: {

          0: {color: 'magenta'},

          //1: {color: 'green'}

          //2: {color: 'blue'},

        }});

        print(histogramfIndex);

        //var WRI.rename('CL');

        Map.addLayer(addWri, {}, 'WRI+NDWI+MNDWI');

        Map.addLayer(fIndex, {}, 'NDWI+MNDWI');

        Map.addLayer(NDWI, {}, 'NDWI');

        Map.addLayer(MNDWI, {}, 'MNDWI');

        Map.addLayer(WRI, {}, 'WRI');

        return
addWri.rename('CL');//.addBands(ee.Image(months).rename('time'));

      };


      var SWEMod = function(S2_image, region){

        var date = ee.Date(S2_image.get('system:time_start'));

        var months = date.difference(ee.Date('1970-01-01'), 'year');
```

```
        var WRI = S2_image.select('B3').add(S2_image.select('B4'))

.divide(S2_image.select('B8').add(S2_image.select('B12'))).rename('WRI')
;//the final index

        var NDWI =  S2_image.normalizedDifference(['B3',
'B8']).rename('NDWI'); //McFeeters

        var MNDWI =  S2_image.normalizedDifference(['B3',
'B12']).rename('MNDWI'); //Xu but with the Swir not MIR (also present on
SNAP);

        var sintIm = ee.Image([WRI, NDWI, MNDWI]);

        return sintIm;

      };


      // Return the DN that maximizes interclass variance in
WRI+NDWI+MNDWI (in the region).

      var otsu = function otsu(histogram) {

          histogram = ee.Dictionary(histogram);


          var counts = ee.Array(histogram.get('histogram'));

          var means = ee.Array(histogram.get('bucketMeans'));

          var size = means.length().get([0]);

          var total = counts.reduce(ee.Reducer.sum(), [0]).get([0]);

          var sum = means.multiply(counts).reduce(ee.Reducer.sum(),
[0]).get([0]);

          var mean = sum.divide(total);


          var indices = ee.List.sequence(1, size);


          // Compute between sum of squares, where each mean partitions
the data.

          var bss = indices.map(function (i) {

              var aCounts = counts.slice(0, 0, i);
```

```
                var aCount = aCounts.reduce(ee.Reducer.sum(),
[0]).get([0]);

                var aMeans = means.slice(0, 0, i);

                var aMean =
aMeans.multiply(aCounts).reduce(ee.Reducer.sum(),
[0]).get([0]).divide(aCount);

                var bCount = total.subtract(aCount);

                var bMean =
sum.subtract(aCount.multiply(aMean)).divide(bCount);

                return
aCount.multiply(aMean.subtract(mean).pow(2)).add(bCount.multiply(bMean.s
ubtract(mean).pow(2)));

            });


            // Return the mean value corresponding to the maximum BSS.

            return means.sort(bss).get([-1]);

        };


        //Quality assessment of the binary classification

        var qualityAssess = function(gTruth, classification, th, area,
id){

          var assessSt = classification.gt(th).multiply(3);

          var assess = assessSt.add(gTruth);

          Map.addLayer(assess, {}, 'Validation Result' + id);

          var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

          var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

          var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

          print(resultEdge, id);

          print(gtSumm, 'Ground Truth');

          print(classSumm, 'Classification');
```

```
  return assess;
};


/*
THIS PART CALLS THE FUNCTIONS DEFINED PREVIOUSLY AND APPLIES THE
FOR EACH AREA
THE CODE NEEDS TO BE UNCOMMENTED FOR EACH AREA OF STUDY
*/


/*
// Case Study Beruwala [1]
Map.addLayer(S2_Beruwala, imageVisParam, 'Sint Band');
//Map.addLayer(S2_Beruwala.normalizedDifference(['B8',
'B4']).gt(0.5), {}, 'moist vegetation');
var WRI_NDWI = SWE(S2_Beruwala, geometryBeruwala);
//var NDVI = S2_Beruwala.normalizedDifference(['B8',
'B4']).lt(0.5);
Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');
//Map.addLayer(regionOfInt,{}, 'Buffer Area');
var o = computeThresholdUsingOtsu(WRI_NDWI, 30, regionOfInt, 1, 1,
false, true, -0.1);


//only otsu on the region of int
var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, scale)).values().get(0));


var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);
thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)
```

```
      var qCB = qualityAssess(Beruwala_GroudTruth, WRI_NDWI,
o.threshold, geometryBeruwala, ' thEDGE');

      var qC = qualityAssess(Beruwala_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryBeruwala, ' thNoEDGE');


      print(o.threshold, ' thEDGE');

      print(thresholdNoEdge, ' thNoEDGE');

      */

      /*

      //Case Study Beruwala modificato[1]

      var thFunction = function(image, bandName){

        var hist =
ee.Dictionary(ee.Dictionary(image.select(bandName).reduceRegion(ee.Reduc
er.histogram(100), regionOfInt, 30)).values().get(0));

        var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

        var bandTh = image.select(bandName);

        return bandTh.gt(ee.Number(thresholdNoEdge)).eq(1);

      };


      Map.addLayer(S2_Beruwala, imageVisParam, 'Sint Band Beruwala');

      Map.addLayer(Beruwala_GroudTruth, gTru, 'G Truth');

      var WRI_NDWI = SWEMod(S2_Beruwala, geometryBeruwala);

      Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');


      //only otsu on the region of int

      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));

      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)
```

```
      //print(o.threshold, 'thEDGE');

      print(thresholdNoEdge, 'Threshold without cannny edge detection');

      //var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');

      var qC = qualityAssess(Beruwala_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryBeruwala, ' no edge');


      var MNDWIth = thFunction(WRI_NDWI, 'MNDWI');

      var NDWIth = thFunction(WRI_NDWI, 'NDWI');

      var WRIth = thFunction(WRI_NDWI, 'WRI');

      var SintImage = MNDWIth.add(NDWIth).add(WRIth).divide(3);

      var SintImageSum = MNDWIth.add(NDWIth).add(WRIth);

      Map.addLayer(SintImageSum, {}, 'somma delle tre immagini');

      var qualityAssessBinClass = function(gTruth, result, area){

        var assessSt = result.eq(3).multiply(3);

        var assess =
assessSt.add(gTruth).add(S2_Beruwala.select(['B3']).lt(2000).eq(4));

        Map.addLayer(assess, validationVisParam, 'Validation Result
sum');

        var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        print(resultEdge, 'Error Matrix');

        //print(gtSumm, 'Ground Truth');

        //print(classSumm, 'Bin classClassification');

        return assess;

      };
```

```
      var qualityBinClass = qualityAssessBinClass(Beruwala_GroudTruth,
SintImageSum, geometryBeruwala);

      //Map.addLayer(SintImage, {}, 'somma degli indici th +
normalizzatione')

      Map.addLayer(SintImageSum, {}, 'somma degli indici th')


      Export.image.toDrive({

        //image: S2_Beruwala.select(['B2', 'B3', 'B4', 'B8', 'B12']),

        image: qualityBinClass,

        description: 'S2_Beruwala_classification',

        fileFormat: 'GeoTIFF',

        scale: 10,

        region: geometryBeruwala

      });

      */


      /*

      // Case Study Viadana [2]

      Map.addLayer(S2_Viadana, imageVisParam, 'Sint Band Viadana');

      Map.addLayer(Viadana_GroudTruth, gTru, 'G Truth');

      var WRI_NDWI = SWE(S2_Viadana, geometryViadana);

      Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');

      //Map.addLayer(regionOfInt,{}, 'Buffer Area');

      var o = computeThresholdUsingOtsu(WRI_NDWI, 30, regionOfInt, 1, 1,
false, true, -0.1);


      //only otsu on the region of int

      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, scale)).values().get(0));
```

```
      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)


      var qCB = qualityAssess(Viadana_GroudTruth, WRI_NDWI, o.threshold,
geometryViadana, ' edge');

      var qC = qualityAssess(Viadana_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryViadana, ' no edge');

      print(o.threshold, 'thEDGE');

      print(thresholdNoEdge, 'thNoEDGE');

      */

      /*

      //Case Study Viadana modificato[2]

      var thFunction = function(image, bandName){

        var hist =
ee.Dictionary(ee.Dictionary(image.select(bandName).reduceRegion(ee.Reduc
er.histogram(100), regionOfInt, 30)).values().get(0));

        var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

        var bandTh = image.select(bandName);

        return bandTh.gt(ee.Number(thresholdNoEdge)).eq(1);

      };


      Map.addLayer(S2_Viadana, imageVisParam, 'Sint Band Viadana');

      Map.addLayer(Viadana_GroudTruth, gTru, 'G Truth');

      var WRI_NDWI = SWEMod(S2_Viadana, geometryViadana);

      Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');


      //only otsu on the region of int

      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));
```

```
      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)

      //print(o.threshold, 'thEDGE');

      print(thresholdNoEdge, 'Threshold without cannny edge detection');

      //var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');

      var qC = qualityAssess(Viadana_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryViadana, ' no edge');


      var MNDWIth = thFunction(WRI_NDWI, 'MNDWI');

      var NDWIth = thFunction(WRI_NDWI, 'NDWI');

      var WRIth = thFunction(WRI_NDWI, 'WRI');

      var SintImage = MNDWIth.add(NDWIth).add(WRIth).divide(3);

      var SintImageSum = MNDWIth.add(NDWIth).add(WRIth);

      Map.addLayer(SintImageSum, {}, 'somma delle tre immagini');

      var qualityAssessBinClass = function(gTruth, SintImageSum, area){

        var assessSt = SintImageSum.eq(3).multiply(3);

        var assess =
assessSt.add(gTruth).add(S2_Viadana.select(['B3']).lt(2000).eq(4));

        Map.addLayer(assess, validationVisParam, 'Validation Result
sum');

        var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        print(resultEdge, 'Error Matrix');

        //print(gtSumm, 'Ground Truth');

        //print(classSumm, 'Bin classClassification');

        return assess;
```

```
};


var qualityBinClass = qualityAssessBinClass(Viadana_GroudTruth,
SintImageSum, geometryViadana);

//Map.addLayer(SintImage, {}, 'somma degli indici th +
normalizzatione')

Map.addLayer(SintImageSum, {}, 'somma degli indici th')

Export.image.toDrive({

  image: qualityBinClass,

  description: 'S2_Viadana_classification',

  fileFormat: 'GeoTIFF',

  scale: 10,

  region: geometryViadana

});

*/


/*

// Case Study Karunki [3]

Map.addLayer(S2_Karunki, imageVisParam, 'Sint Band Karunki');

Map.addLayer(Karunki_GroudTruth, gTru, 'G Truth');

var WRI_NDWI = SWE(S2_Karunki, geometryKarunki);

Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');

/*

//Map.addLayer(regionOfInt,{}, 'Buffer Area');

//var S2_clip = S2_Karunki.select(['B1', 'B2', 'B3', 'B4', 'B5',
'B6', 'B7', 'B8', 'B9', 'B10', 'B11', 'B12']).clip(geometryKarunki);

//var o = computeThresholdUsingOtsu(WRI_NDWI.select('CL'), 30,
regionOfInt, 1, 1, false, true, -0.1);


//only otsu on the region of int
```

```
      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));


      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)

      //print(o.threshold, 'thEDGE');

      print(thresholdNoEdge, 'Threshold without cannny edge detection');

      //var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');

      var qC = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryKarunki, ' no edge');

      //Map.addLayer(WRI_NDWI.gt(0.719), {}, 'test')

      */

      /*

      // Case Study Karunki SWE modificato [3]

      Map.addLayer(S2_Karunki, imageVisParam, 'Sint Band Karunki');

      Map.addLayer(Karunki_GroudTruth, gTru, 'G Truth');

      var WRI_NDWI = SWE(S2_Karunki, geometryKarunki);

      var thFunction = function(image, bandName){

        var hist =
ee.Dictionary(ee.Dictionary(image.select(bandName).reduceRegion(ee.Reduc
er.histogram(100), regionOfInt, 30)).values().get(0));

        var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

        var bandTh = image.select(bandName);

        return bandTh.gt(ee.Number(thresholdNoEdge)).eq(1);

      };


      Map.addLayer(S2_Karunki, imageVisParam, 'Sint Band Karunki');

      Map.addLayer(Karunki_GroudTruth, gTru, 'G Truth');
```

```
var WRI_NDWI = SWEMod(S2_Karunki, geometryKarunki);

Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');


//only otsu on the region of int

var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));

var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)

//print(o.threshold, 'thEDGE');

print(thresholdNoEdge, 'Threshold without cannny edge detection');

//var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');

var qC = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryKarunki, ' no edge');


var MNDWIth = thFunction(WRI_NDWI, 'MNDWI');

var NDWIth = thFunction(WRI_NDWI, 'NDWI');

var WRIth = thFunction(WRI_NDWI, 'WRI');

var SintImage = MNDWIth.add(NDWIth).add(WRIth).divide(3);

var SintImageSum = MNDWIth.add(NDWIth).add(WRIth);

Map.addLayer(SintImageSum, {}, 'somma delle tre immagini');

var qualityAssessBinClass = function(gTruth, SintImageSum, area){

  var assessSt = SintImageSum.eq(3).multiply(3);

  var assess =
assessSt.add(gTruth).add(S2_Karunki.select(['B3']).lt(2000).eq(4));

  Map.addLayer(assess, validationVisParam, 'Validation Result
sum');

  var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

  var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);
```

```
        var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);
        print(resultEdge, 'Error Matrix');
        //print(gtSumm, 'Ground Truth');
        //print(classSumm, 'Bin classClassification');
        return assess;
      };

      var qualityBinClass = qualityAssessBinClass(Karunki_GroudTruth,
SintImageSum, geometryKarunki);
      Map.addLayer(SintImage, {}, 'somma degli indici th +
normalizzatione')
      Map.addLayer(SintImageSum, {}, 'somma degli indici th')
      Export.image.toDrive({
        image: qualityBinClass,
        description: 'S2_Karunki_classification',
        fileFormat: 'GeoTIFF',
        scale: 10,
        region: geometryKarunki
      });
      */

      //Case study validation and assessment
      /*
      // Case Study Puerto [4]
      Map.addLayer(S2_Puerto, imageVisParam, 'Sint Band Puerto');
      Map.addLayer(Puerto_GroudTruth, gTru, 'G Truth');
      var WRI_NDWI = SWE(S2_Puerto);
      Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');
      //Map.addLayer(regionOfInt,{}, 'Buffer Area');
```

```
      var o = computeThresholdUsingOtsu(WRI_NDWI, 30, regionOfInt, 1, 1,
false, true, -0.1);


      //only otsu on the region of int

      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, scale)).values().get(0));


      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)


      var qCB = qualityAssess(Puerto_GroudTruth, WRI_NDWI, o.threshold,
geometryPuerto, ' edge');

      var qC = qualityAssess(Puerto_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryPuerto, ' no edge');

      print(o.threshold, 'thEDGE');

      print(thresholdNoEdge, 'thNoEDGE');

      */



      // Case Study Puerto SWE modificato [4]

      var thFunction = function(image, bandName){

        var hist =
ee.Dictionary(ee.Dictionary(image.select(bandName).reduceRegion(ee.Reduc
er.histogram(100), regionOfInt, 30)).values().get(0));

        var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

        var bandTh = image.select(bandName);

        return bandTh.gt(ee.Number(thresholdNoEdge)).eq(1);

      };


      Map.addLayer(S2_Puerto, imageVisParam, 'Sint Band Puerto');
```

```
Map.addLayer(Puerto_GroudTruth, gTru, 'G Truth');

var WRI_NDWI = SWEMod(S2_Puerto, geometryPuerto);

Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');


//only otsu on the region of int

var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));

var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)

//print(o.threshold, 'thEDGE');

print(thresholdNoEdge, 'Threshold without cannny edge detection');

//var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');

var qC = qualityAssess(Puerto_GroudTruth, WRI_NDWI,
thresholdNoEdge, geometryPuerto, ' no edge');


var MNDWIth = thFunction(WRI_NDWI, 'MNDWI');

var NDWIth = thFunction(WRI_NDWI, 'NDWI');

var WRIth = thFunction(WRI_NDWI, 'WRI');

var SintImage = MNDWIth.add(NDWIth).add(WRIth).divide(3);

var SintImageSum =
MNDWIth.add(NDWIth).add(WRIth)//.mask(S2_Puerto.select(['B3']).lt(2000))
//;

Map.addLayer(SintImageSum, {}, 'somma delle tre immagini');

var qualityAssessBinClass = function(gTruth, result, area){

  var assessSt = result.eq(3).multiply(3);

  var assess =
assessSt.add(gTruth).add(S2_Puerto.select(['B3']).lt(3500).eq(4));

  //Map.addLayer(assess,
{"opacity":1,"bands":["WRI"],"max":4,"palette":["19ff15","ffa434","ff000
0","0726ff"]}, 'Validation Result sum');
```

```
        Map.addLayer(assess, validationVisParam, 'Validation Result
sum');

        var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);

        print(resultEdge, 'Error Matrix');

        //print(gtSumm, 'Ground Truth');

        //print(classSumm, 'Bin classClassification');

        return assess;

      };


      var qualityBinClass = qualityAssessBinClass(Puerto_GroudTruth,
SintImageSum, geometryPuerto);

      Map.addLayer(SintImage, {}, 'somma degli indici th +
normalizzatione')

      Map.addLayer(SintImageSum, {}, 'somma degli indici th')

      //Puerto

      Export.image.toDrive({

        //image: S2_Puerto.select(['B2', 'B3', 'B4', 'B8', 'B12']),

        image: qualityBinClass,

        description: 'S2_Puerto_classification',

        fileFormat: 'GeoTIFF',

        scale: 10,

        region: geometryPuerto

      });


      /*

      // Case Study Magwitz [5]
```

```
Map.addLayer(S2_Mag, imageVisParam, 'Sint Band Mag');

Map.addLayer(Mag_GroudTruth, gTru, 'G Truth');

var WRI_NDWI = SWE(S2_Mag, geometryMag);

var huggel = S2_Mag.normalizedDifference(['B2', 'B8'])

Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');

Map.addLayer(huggel, {}, 'huggel');

//Map.addLayer(regionOfInt,{}, 'Buffer Area');

var o = computeThresholdUsingOtsu(WRI_NDWI, 30, regionOfInt, 1, 1,
false, true, -0.1);


//only otsu on the region of int

var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, scale)).values().get(0));


var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);

thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)


var qCB = qualityAssess(Mag_GroudTruth, WRI_NDWI, o.threshold,
geometryMag, ' edge');

var qC = qualityAssess(Mag_GroudTruth, WRI_NDWI, thresholdNoEdge,
geometryMag, ' no edge');

print(o.threshold, 'thEDGE');

print(thresholdNoEdge, 'thNoEDGE');

*/

/*

// Case Study Magwitz modificato [5]

var thFunction = function(image, bandName){

  var hist =
ee.Dictionary(ee.Dictionary(image.select(bandName).reduceRegion(ee.Reduc
er.histogram(100), regionOfInt, 30)).values().get(0));
```

```
        var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);
        var bandTh = image.select(bandName);
        return bandTh.gt(ee.Number(thresholdNoEdge)).eq(1);
      };


      Map.addLayer(S2_Mag, imageVisParam, 'Sint Band Mag');
      Map.addLayer(Mag_GroudTruth, gTru, 'G Truth');
      var WRI_NDWI = SWEMod(S2_Mag, geometryMag);
      Map.addLayer(WRI_NDWI, {}, 'Original data no th applied');


      //only otsu on the region of int
      var hist =
ee.Dictionary(ee.Dictionary(WRI_NDWI.reduceRegion(ee.Reducer.histogram(1
00), regionOfInt, 30)).values().get(0));
      var thresholdNoEdge =
ee.Algorithms.If(hist.contains('bucketMeans'), otsu(hist), 0);
      thresholdNoEdge = ee.Number(thresholdNoEdge); //.add(0.05)
      //print(o.threshold, 'thEDGE');
      print(thresholdNoEdge, 'Threshold without cannny edge detection');
      //var qCB = qualityAssess(Karunki_GroudTruth, WRI_NDWI,
o.threshold, geometryKarunki, ' edge');
      var qC = qualityAssess(Mag_GroudTruth, WRI_NDWI, thresholdNoEdge,
geometryMag, ' no edge');


      var MNDWIth = thFunction(WRI_NDWI, 'MNDWI');
      var NDWIth = thFunction(WRI_NDWI, 'NDWI');
      var WRIth = thFunction(WRI_NDWI, 'WRI');
      var SintImage = MNDWIth.add(NDWIth).add(WRIth).divide(3);
      var SintImageSum =
MNDWIth.add(NDWIth).add(WRIth).mask(S2_Mag.select(['B3']).lt(2000));//;
      Map.addLayer(SintImageSum, {}, 'somma delle tre immagini');
```

```
      var qualityAssessBinClass = function(gTruth, result, area){
        var assessSt = result.eq(3).multiply(3);
        var assess =
assessSt.add(gTruth)//.add(S2_Mag.select(['B3']).lt(3500).eq(3));
        Map.addLayer(assess, validationVisParam, 'Validation Result
sum');
        var resultEdge =
assess.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);
        var gtSumm =
gTruth.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);
        var classSumm =
assessSt.reduceRegion(ee.Reducer.frequencyHistogram(), area, 10);
        print(resultEdge, 'Error Matrix');
        //print(gtSumm, 'Ground Truth');
        //print(classSumm, 'Bin classClassification');
        return assess;
      };
      var qualityBinClass = qualityAssessBinClass(Mag_GroudTruth,
SintImageSum, geometryMag);
      Map.addLayer(SintImage, {}, 'somma degli indici th +
normalizzatione')
      Map.addLayer(SintImageSum, {}, 'somma degli indici th')
      */


      /*This function allows to export the results*/
      //Magwitz
      Export.image.toDrive({
        //image: S2_Mag.select(['B2', 'B3', 'B4', 'B8', 'B12']),
        image: qualityBinClass,
        description: 'S2_Magwitz_classification',
        fileFormat: 'GeoTIFF',
        scale: 10,
```

```
        region: geometryMag
      });
```

2_Code used for the production process

```
  var S2 = ee.ImageCollection("COPERNICUS/S2"),
      imageVisParam =
{"opacity":1,"bands":["B4","B3","B2"],"min":828.1,"max":2004.9,"gamma":1
},
      falseColorVisParam =
{"opacity":1,"bands":["B8","B4","B3"],"min":917.3,"max":4463.7,"gamma":1
},
      geometry2 = /* color: #d63000 */ee.Geometry.MultiPoint(
          [[7.232224081009804, 45.009918149223104],
           [7.7213584224881515, 45.08506671069958],
           [7.721894369577967, 45.05022751285707],
           [7.674387522786446, 45.06157011410555],
           [91.63980936330586, 26.17751179984684],
           [91.62040364599443, 26.20809583659951]]),
      table = ee.FeatureCollection("users/consandu/scenario"),
      geometry3 = /* color: #d9f116
*/ee.Geometry.Point([91.55054544729023, 26.149316446621473]),
      samplingPoint = /* color: #2f44d6
*/ee.Geometry.Point([24.701388842320284, 43.81301881361876]);
  var AOI = table;
  Map.addLayer(AOI, {},'AOI');
  var start = ee.Date('2017-01-01');
  var finish = ee.Date(Date.now());
```

```
   //var start = ee.Date('2017-11-01');
   //var finish = ee.Date('2018-01-14');
   var regionOfInt = samplingPoint.buffer(2000);

   var filteredS2Collection = S2.filterBounds(samplingPoint)
                               .filterDate(start, finish)
.filter(ee.Filter.lt('CLOUDY_PIXEL_PERCENTAGE', 20));

   // Return the DN that maximizes interclass variance in each of the
mask (in the region).
   var otsu = function otsu(histogram) {
       histogram = ee.Dictionary(histogram);

       var counts = ee.Array(histogram.get('histogram'));
       var means = ee.Array(histogram.get('bucketMeans'));
       var size = means.length().get([0]);
       var total = counts.reduce(ee.Reducer.sum(), [0]).get([0]);
       var sum = means.multiply(counts).reduce(ee.Reducer.sum(),
[0]).get([0]);
       var mean = sum.divide(total);

       var indices = ee.List.sequence(1, size);
```

```
      // Compute between sum of squares, where each mean partitions the
data.
      var bss = indices.map(function (i) {
          var aCounts = counts.slice(0, 0, i);
          var aCount = aCounts.reduce(ee.Reducer.sum(), [0]).get([0]);
          var aMeans = means.slice(0, 0, i);
          var aMean = aMeans.multiply(aCounts).reduce(ee.Reducer.sum(),
[0]).get([0]).divide(aCount);
          var bCount = total.subtract(aCount);
          var bMean =
sum.subtract(aCount.multiply(aMean)).divide(bCount);
          return
aCount.multiply(aMean.subtract(mean).pow(2)).add(bCount.multiply(bMean.s
ubtract(mean).pow(2)));
      });

      // Return the mean value corresponding to the maximum BSS.
      return means.sort(bss).get([-1]);
  };

  var watFrequency = function(S2WatmaskCollection){
    var NumWatObs = S2WatmaskCollection.select('ND').sum();
    //print(NumWatObs, 'n obs');
    var NumWatObsValid = S2WatmaskCollection.select('ND').count();
    //print(NumWatObsValid, 'v obs');
    var NumWatObsFrequency = NumWatObs.divide(NumWatObsValid);
    return NumWatObsFrequency;
```

```
};


//Create filter
var filter = ee.Filter.equals({
  leftField: 'system:index',
  rightField: 'system:index'
});
var filterWRI = ee.Filter.equals({
  leftField: 'secondId',
  rightField: 'system:index'
});


//Create Simple Join
var innerJoin = ee.Join.inner();


var computeMNDWI = function(image){
  var doy = image.date().getRelative('day', 'year');
  var doyBand = ee.Image.constant(doy).uint16().rename('doy');
  var mndwi = image.normalizedDifference(['B3',
'B12']).rename('MNDWI');
  //remember to update time band
  var hist =
ee.Dictionary(ee.Dictionary(mndwi.reduceRegion(ee.Reducer.histogram(100)
, regionOfInt, 30)).values().get(0));
  var thresholdNoEdge = ee.Algorithms.If(hist.contains('bucketMeans'),
otsu(hist), 0);
```

```
    thresholdNoEdge = ee.Number(thresholdNoEdge);

    return
mndwi.gt(thresholdNoEdge).eq(1).mask(image.select(['B3']).lt(2000)).addB
ands(doyBand).set({secondId: image.get('system:index')});

  };


  var computeNDWI = function(image){

    var doy = image.date().getRelative('day', 'year');

    var doyBand = ee.Image.constant(doy).uint16().rename('doy');

    var ndwi = image.normalizedDifference(['B3',
'B8']).rename('NDWI')//.addBands(date);

    var hist =
ee.Dictionary(ee.Dictionary(ndwi.reduceRegion(ee.Reducer.histogram(100),
regionOfInt, 30)).values().get(0));

    var thresholdNoEdge = ee.Algorithms.If(hist.contains('bucketMeans'),
otsu(hist), 0);

    thresholdNoEdge = ee.Number(thresholdNoEdge);

    return
ndwi.gt(thresholdNoEdge).eq(1).addBands(doyBand).mask(image.select(['B3'
]).lt(2000)).addBands(doyBand);

  };


  var computeWRI = function(image){

    var doy = image.date().getRelative('day', 'year');

    var doyBand = ee.Image.constant(doy).uint16().rename('doy');

    var wri =
image.select('B3').add(image.select('B4')).divide(image.select('B8').add
(image.select('B12'))).rename('WRI');
```

```
    var hist =
ee.Dictionary(ee.Dictionary(wri.reduceRegion(ee.Reducer.histogram(100),
regionOfInt, 30)).values().get(0));

    var thresholdNoEdge = ee.Algorithms.If(hist.contains('bucketMeans'),
otsu(hist), 0);

    thresholdNoEdge = ee.Number(thresholdNoEdge);

    return
wri.gt(thresholdNoEdge).eq(1).mask(image.select(['B3']).lt(2000)).addBan
ds(doyBand);

  };



  var mergeBands = function(aRow) {

    var anImage = ee.Image.cat(aRow.get('primary'),
aRow.get('secondary'));

    return anImage;

  };



  var sumOfThBands = function(image){

    var out =
image.select('MNDWI').add(image.select('NDWI')).add(image.select('WRI'))
.rename('ND');

    return out.eq(3).addBands(image.select("doy"));

  };



  var sumOfThBandsNumObs = function(image){

    var out =
image.select('MNDWI').add(image.select('NDWI')).add(image.select('WRI'))
.rename('ND');

    return out.eq(3).multiply(100).addBands(image.select("doy"));
```

```
};


/*this is the computation of all the NDVI of the period June to
September
 of 2018*/



var mndwiCollection = filteredS2Collection.map(computeMNDWI);
var ndwiCollection = filteredS2Collection.map(computeNDWI);
var wriCollection = filteredS2Collection.map(computeWRI);




/* compute the water frequency
*/


var ndwiMndwiJoin = innerJoin.apply(mndwiCollection, ndwiCollection,
filter);
var mergedMndwiNdwi = ndwiMndwiJoin.map(mergeBands)
var ndwiMndwiWriJoin = innerJoin.apply(mergedMndwiNdwi, wriCollection,
filterWRI);
var mergedNdwiMndwiWRI = ndwiMndwiWriJoin.map(mergeBands)
//print(mergedNdwiMndwiWRI)
var mergedSumedNdwiMndwiWRI =
ee.ImageCollection(mergedNdwiMndwiWRI).map(sumOfThBands);
var mergedSumedNdwiMndwiWRInumObs =
ee.ImageCollection(mergedNdwiMndwiWRI).map(sumOfThBandsNumObs);
```

```
//print(mergedSumedNdwiMndwiWRI);

var waterFrequency = watFrequency(mergedSumedNdwiMndwiWRI);


//var linearFit = ndwiCollection.reduce(ee.Reducer.linearFit());

//Map.addLayer(linearFit.select(['scale']), {min: -90, max: 90,
palette: ['white','#0088ff', 'navy']}, 'Linear Fit');

//this is only for visualize purpose

//var leastCloudyS2 = filteredS2Collection.first();

//Map.addLayer(leastCloudyS2, falseColorVisParam, 'cape town s2');

Map.addLayer(mergedSumedNdwiMndwiWRInumObs, {}, 'separate images of
reelaborated collection', false);

Map.addLayer(waterFrequency, {min: 0, max: 1, palette:
['white','#0088ff', 'navy']}, 'Water Occurance');

//Map.addLayer(waterFrequency, {}, 'Water frequency', false);


/*

// Export the image, specifying scale and region.

Export.image.toDrive({

  image: waterFrequency,

  description: 'Permanent Water',

  scale: 30,

  region: AOI

});

*/
```